

PATENT
81940.0050
Express Mail Label No. EV 325 216 955 US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:	Art Unit: Not assigned
Sigeo HOMMA et al.	Examiner: Not assigned
Serial No: Not assigned	
Filed: July 25, 2003	
For: STORAGE APPARATUS AND ITS MANAGEMENT METHOD	

TRANSMITTAL OF PRIORITY DOCUMENT

Mail Stop PATENT APPLICATION
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

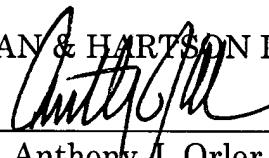
Dear Sir:

Enclosed herewith is a certified copy of Japanese patent application No. 2003-048483 filed February 26, 2003, from which priority is claimed under 35 U.S.C. § 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Respectfully submitted,

HOGAN & HARTSON L.L.P.

Date: July 25, 2003
By: 
Anthony J. Orler
Registration No. 41,232
Attorney for Applicant(s)

500 South Grand Avenue, Suite 1900
Los Angeles, California 90071
Telephone: 213-337-6700
Facsimile: 213-337-6701

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application:

2003年 2月26日

出願番号

Application Number:

特願2003-048483

[ST.10/C]:

[JP2003-048483]

出願人

Applicant(s):

株式会社日立製作所

2003年 6月 9日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3044539

【書類名】 特許願
【整理番号】 K02017521A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 3/06
【発明者】
【住所又は居所】 神奈川県小田原市中里322番地2号 株式会社日立製作所 R A I D システム事業部内
【氏名】 本間 繁雄
【発明者】
【住所又は居所】 神奈川県小田原市中里322番地2号 株式会社日立製作所 R A I D システム事業部内
【氏名】 渡辺 治明
【発明者】
【住所又は居所】 神奈川県川崎市麻生区玉禅寺1099番地 株式会社日立製作所 システム開発研究所内
【氏名】 水野 陽一
【発明者】
【住所又は居所】 神奈川県川崎市麻生区玉禅寺1099番地 株式会社日立製作所 システム開発研究所内
【氏名】 田中 淳
【特許出願人】
【識別番号】 000005108
【氏名又は名称】 株式会社日立製作所
【代理人】
【識別番号】 100075096
【弁理士】
【氏名又は名称】 作田 康夫
【手数料の表示】
【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】ストレージ装置およびその管理方法

【特許請求の範囲】

【請求項1】

コンピュータからのコマンドに応答して動作するストレージ装置であって、記憶領域と、前記記憶領域内の第1の記憶領域に前記記憶領域内の第2の記憶領域を対応付け、前記第2の記憶領域のデータに対するRead命令に応答して、前記第1の記憶領域から前記第2の記憶領域への前記データの複写を許容するように制御するための第1のコマンドと、前記第1の記憶領域と前記第2の記憶領域との前記対応付けの解消を制御するための第2のコマンドとに応答して動作する制御装置とを含むストレージ装置。

【請求項2】

請求項1のストレージ装置において、前記データの複写を制御する前記制御装置は、前記第1の記憶領域からデータを読み出し、該読み出したデータを前記コンピュータへ出力した後に、前記読み出したデータを前記第2の記憶領域への書き込むように制御するストレージ装置。

【請求項3】

請求項1のストレージ装置において、前記制御装置は、前記第1のコマンドに応答して、前記第1の記憶領域から前記第2の記憶領域へ複写したデータがある場合、該データを無意にするように制御するストレージ装置。

【請求項4】

請求項1のストレージ装置において、前記第1のコマンドは、さらに前記第2の記憶領域に対するWrite命令に応答して、前記第1の記憶領域から前記第2の記憶領域へのデータの複写を許容するように制御するためのコマンドであるストレージ装置。

【請求項5】

請求項4のストレージ装置において、前記制御装置は、前記第2の記憶領域へ複写されたデータの上に、前記Write命令で指示されたデータを書き込むように制御するストレージ装置。

【請求項6】

請求項5のストレージ装置において、前記制御装置は、前記第2のコマンドに応答して、前記第2の記憶領域に書き込まれた、前記Write命令で指示されたデータの前記第1の記憶領域への複写を制御するストレージ装置。

【請求項7】

請求項4のストレージ装置において、前記制御装置は、前記第2のコマンドに応答して、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からのデータの複写を制御するストレージ装置。

【請求項8】

請求項1のストレージ装置において、前記制御装置は、前記第2のコマンドに応答して、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からのデータの複写を制御するストレージ装置。

【請求項9】

請求項1のストレージ装置において、前記制御装置は、前記第1のコマンドに応答して、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理するための管理情報を生成し、前記第2のコマンドに応答して、前記管理情報を解放するストレージ装置。

【請求項10】

請求項1のストレージ装置において、前記制御装置は、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からのデータの複写と、前記第1の記憶領域と前記第2の記憶領域との前記対応付けの解消を制御するための第3のコマンドとに応答する制御装置とを含むストレージ装置。

【請求項11】

請求項10のストレージ装置において、前記制御装置は、前記コンピュータから選択的に発行される前記第2のコマンドと前記第3のコマンドとに応答するストレージ装置。

【請求項12】

コンピュータからのコマンドに応答して動作するストレージ装置であって、記憶領域と、前記記憶領域内の第1の記憶領域に前記記憶領域内の第2の記憶領域を対応付け、前記第2の記憶領域のデータに対するRead命令に応答した、前記第1の記憶領域から前記第2の記憶領域への前記データの複写と、前記第2の記憶領域に対するWrite命令に応答して、前記第2の記憶領域へ前記Write命令で指示されたデータの書き込みを許容するように制御するための第1のコマンドに応答して動作し、前記第2の記憶領域に前記書き込まれたデータの前記第1の記憶領域への複写を制御する制御装置とを含むストレージ装置。

【請求項13】

請求項12のストレージ装置において、前記制御装置は、Read命令に応答した前記データの複写を、前記第1の記憶領域から前記データを読み出し、該読み出したデータを前記コンピュータの出力した後に、前記読み出したデータを前記第2の記憶領域への書き込むように制御するストレージ装置。

【請求項14】

請求項12のストレージ装置において、前記制御装置は、前記第1の記憶領域への前記書き込まれたデータの複写を第2のコマンドとに応答して制御するする制御装置であるストレージ装置。

【請求項15】

請求項12のストレージ装置において、前記制御装置は、前記第1のコマンドに応答して、前記Write命令で指示されたデータの書き込みの前に、前記Write命令で指示された前記第2の記憶領域に前記第1の記憶領域のデータの複写を制御するストレージ装置。

【請求項16】

請求項12のストレージ装置において、前記制御装置は、前記第1のコマンドに応答して、前記第1の記憶領域から前記第2の記憶領域へ複写したデータがある場合、該データを無意にするように制御するストレージ装置。

【請求項17】

請求項12のストレージ装置において、前記制御装置は、前記第1の記憶領域

と前記第2の記憶領域との前記対応付けの解消を制御するための第3のコマンドに応答するストレージ装置。

【請求項18】

請求項17のストレージ装置において、前記制御装置は、前記第1のコマンドに応答して、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理するための管理情報を生成し、前記第3のコマンドに応答して、前記管理情報を解放するストレージ装置。

【請求項19】

ストレージ装置であって、記憶領域と、前記記憶領域内の第1の記憶領域に前記記憶領域内の第2の記憶領域を対応付け、前記第2の記憶領域のデータに対するRead命令に応答して、前記第1の記憶領域から前記第2の記憶領域への前記データの複写を許容する第1の制御状態と、前記第1の記憶領域と前記第2の記憶領域との前記対応付けを解消した第2の制御状態との遷移を制御する制御装置とを含むストレージ装置。

【請求項20】

請求項19のストレージ装置において、前記データの複写を制御する前記制御装置は、前記第1の記憶領域からデータを読み出し、該読み出したデータを前記コンピュータへ出力した後に、前記読み出したデータを前記第2の記憶領域への書き込むように制御するストレージ装置。

【請求項21】

請求項19のストレージ装置において、前記制御装置は、前記第1の制御状態で、前記第2の記憶領域に対するWrite命令に応答して、前記第1の記憶領域から前記第2の記憶領域へのデータの複写を許容するように制御するストレージ装置。

【請求項22】

請求項19のストレージ装置において、前記制御装置は、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からのデータの複写を制御するストレージ装置。

【請求項23】

請求項19のストレージ装置において、前記制御装置は、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第1の記憶領域に前記第2の記憶領域からのデータが複写されていない領域に、前記第2の記憶領域からのデータを複写するストレージ装置。

【請求項24】

請求項19のストレージ装置において、前記制御装置は、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理するための管理情報を生成し、前記第2の制御状態から前記第1の制御状態へ遷移させるときに、前記管理情報を解放するストレージ装置。

【請求項25】

記憶領域を有するストレージ装置の管理方法であって、コンピュータからの第1のコマンドに応答して、前記記憶領域内の第1の記憶領域に前記記憶領域内の第2の記憶領域を対応付け、前記第2の記憶領域のデータに対する、コンピュータからのRead命令に応答して、前記第1の記憶領域から前記第2の記憶領域への前記データの複写し、前記コンピュータからの第2のコマンドに応答して、前記第1の記憶領域と前記第2の記憶領域との前記対応付けを解消するストレージ装置の管理方法。

【請求項26】

請求項25のストレージ装置の管理方法において、前記データを複写するステップは、前記第1の記憶領域からデータを読み出し、該読み出したデータを前記コンピュータへ出力し、前記読み出したデータを前記第2の記憶領域への書き込むストレージ装置の管理方法。

【請求項27】

請求項25のストレージ装置の管理方法において、前記第1のコマンドに応答して、前記第1の記憶領域から前記第2の記憶領域へ複写したデータがある場合、該データを無意にするストレージ装置の管理方法。

【請求項28】

請求項27のストレージ装置において、前記第2の記憶領域に対するWrite命令に応答して、前記第1の記憶領域から前記第2の記憶領域へのデータを複写するストレージ装置の管理方法。

【請求項29】

請求項28のストレージ装置において、前記第2の記憶領域へ複写されたデータの上に、前記Write命令で指示されたデータを書き込むストレージ装置の管理方法。

【請求項30】

請求項29のストレージ装置の管理方法において、前記第2のコマンドに応答して、前記第2の記憶領域に書き込まれた、前記Write命令で指示されたデータを前記第1の記憶領域へ複写するストレージ装置の管理方法。

【請求項31】

請求項29のストレージ装置の管理方法において、前記第2のコマンドに応答して、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からデータを複写するストレージ装置の管理方法。

【請求項32】

請求項25のストレージ装置の管理方法において、前記第2のコマンドに応答して、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からデータを複写するストレージ装置の管理方法。

【請求項33】

請求項25のストレージ装置の管理方法において、前記第1のコマンドに応答して、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理するための管理情報を生成し、

前記第2のコマンドに応答して、前記管理情報を解放するストレージ装置の管理方法。

【請求項34】

請求項25のストレージ装置の管理方法において、前記制御装置は、前記第2

の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からデータを複写し、前記第1の記憶領域と前記第2の記憶領域との前記対応付けを解消するストレージ装置の管理方法。

【請求項35】

請求項34のストレージ装置の管理方法において、前記コンピュータから選択的に発行される前記第2のコマンドと前記第3のコマンドとに応答するストレージ装置の管理方法。

【請求項36】

記憶領域を有するストレージ装置の管理方法であって、前記記憶領域内の第1の記憶領域に前記記憶領域内の第2の記憶領域を対応付け、前記第2の記憶領域のデータに対する、コンピュータからのRead命令に応答して、前記第1の記憶領域から前記第2の記憶領域へ前記データを複写し、前記第2の記憶領域に対する、前記コンピュータからのWrite命令に応答して、前記第2の記憶領域へ前記Write命令で指示されたデータの書き込み、前記第2の記憶領域に前記書き込まれたデータを前記第1の記憶領域へ複写するストレージ装置の管理方法。

【請求項37】

請求項36のストレージ装置の管理方法において、前記データを複写するステップは、前記第1の記憶領域から前記データを読み出し、該読み出したデータを前記コンピュータの出力し、前記読み出したデータを前記第2の記憶領域への書き込むストレージ装置の管理方法。

【請求項38】

請求項36のストレージ装置の管理方法において、前記書き込まれたデータの前記第1の記憶領域へ複写するステップは、前記コンピュータからのコマンドとに応答して実行するストレージ装置の管理方法。

【請求項39】

請求項36のストレージ装置の管理方法において、前記Write命令で指示されたデータの書き込みの前に、前記Write命令で指示された前記第2の記憶領域に前記第1の記憶領域のデータを複写するストレージ装置の管理方法。

【請求項40】

請求項36のストレージ装置の管理方法において、前記コンピュータからのコマンドに応答して、前記第1の記憶領域から前記第2の記憶領域へ複写したデータがある場合、該データを無意にするように制御するストレージ装置の管理方法。

【請求項41】

請求項36のストレージ装置の管理方法において、前記コンピュータからのコマンドに応答して、前記第1の記憶領域と前記第2の記憶領域との前記対応付けを解消するストレージ装置の管理方法。

【請求項42】

請求項41のストレージ装置の管理方法において、前記第1の記憶領域に前記第2の記憶領域の対応付けのステップは、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理する管理情報を生成するステップを含み、前記コマンドに応答して、前記管理情報を解放するストレージ装置の管理方法。

【請求項43】

記憶領域を有するストレージ装置の管理方法であって、前記記憶領域内の第1の記憶領域と前記記憶領域内の第2の記憶領域とを対応付けた第1の制御状態と、前記第1の記憶領域と前記第2の記憶領域との前記対応付けを解消した第2の制御状態との遷移を制御し、前記第1の制御状態において、前記第2の記憶領域のデータに対するRead命令に応答して、前記第1の記憶領域から前記第2の記憶領域への前記データを複写するストレージ装置の管理方法。

【請求項44】

請求項43のストレージ装置の管理方法において、前記データを複写するステップは、前記第1の記憶領域からデータを読み出し、該読み出したデータを前記コンピュータへ出力し、前記読み出したデータを前記第2の記憶領域への書き込むストレージ装置の管理方法。

【請求項45】

請求項43のストレージ装置の管理方法において、前記第1の制御状態において、前記第2の記憶領域に対するWrite命令に応答して、前記第1の記憶領域から前記第2の記憶領域へデータを複写するストレージ装置の管理方法。

【請求項4 6】

請求項4 3のストレージ装置の管理方法において、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第2の記憶領域に前記第1の記憶領域からのデータが複写されていない領域に、前記第1の記憶領域からデータを複写するストレージ装置の管理方法。

【請求項4 7】

請求項4 3のストレージ装置の管理方法において、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第1の記憶領域に前記第2の記憶領域からのデータが複写されていない領域に、前記第2の記憶領域からのデータを複写するストレージ装置の管理方法。

【請求項4 8】

請求項4 3のストレージ装置の管理方法において、前記第1の制御状態から前記第2の制御状態へ遷移させるときに、前記第1の記憶領域と前記第2の記憶領域との対応付けを管理するための管理情報を生成し、前記第2の制御状態から前記第1の制御状態へ遷移させるときに、前記管理情報を解放するストレージ装置の管理方法。

【請求項4 9】

ストレージ装置であって、記憶領域と、前記記憶領域内の第1の記憶領域と前記記憶領域内の第2の記憶領域とが対応付けられて制御される第1の状態から前記対応付けを解消して制御される第2の状態への遷移制御するために、前記対応付けを解消する第1の制御モードと前記第1の記憶領域のデータを前記第2の記憶領域に複写した後に、前記対応付けを解消する第2の制御モードを有し、前記第1の制御モードと前記第2の制御モードを選択的に実行する制御装置とを備えたストレージ装置。

【請求項5 0】

ストレージ装置であって、記憶領域と、前記記憶領域内の第1の記憶領域と前記記憶領域内の第2の記憶領域とが対応付けられて制御される第1の状態から前記対応付けを解消して制御される第2の状態への遷移制御するために、前記対応付けを解消する第1の制御モードと前記第2の記憶領域のデータを前記第1の記

憶領域に複写した後に、前記対応付けを解消する第2の制御モードを有し、前記第1の制御モードと前記第2の制御モードを選択的に実行する制御装置とを備えたストレージ装置。

【請求項51】

ストレージ装置であって、記憶領域と、メモリに格納されている管理情報を前記記憶領域内の第1の記憶領域と前記記憶領域内の第2の記憶領域との対応付けを管理するための制御情報、および前記第1の記憶領域へのアクセスに伴うデータの状態と前記第2の記憶領域へのアクセスに伴うデータの状態との対応付けを管理するための制御情報として共通的に用い、前記データが格納されている前記第1の記憶領域及び前記第2の記憶領域を管理する制御装置とを備えたストレージ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ストレージ装置およびその管理方法に関する。

【0002】

【従来の技術】

コンピュータシステムに用いられるストレージ装置に格納されたデータは、装置障害、誤操作などにより喪失したデータを回復するために、磁気テープ等に定期的にバックアップデータとして保存される。バックアップデータは、バックアップのために保存される場合の他に、バッチ処理のデータとして、あるいはアプリケーションプログラムのテストデータ等として用いられる。

【0003】

このようにして用いられるバックアップデータは、ある特定の時点のデータであることが重要であり、それが保証されなければならない。バックアップ処理中に、そのデータの一部が更新されると、データの不整合が生じる。

【0004】

バックアップデータを得る処理は、データのコピー処理が繰り返されるもので長時間をする。データの不整合を生じないようにするために、データが更新

されないようにするためにバックアップ以外の処理を停止すればよいが、オンライン処理を行うシステムでは長時間の処理の中止は許されない。

【0005】

(第1の従来技術) バックアップデータを得る方法として、ミラーリングという技術がある(例えば、特許文献1参照)。これは、ストレージ装置内のボリュームを物理的に二重化する技術である。データの一致のためにコピー処理を行うが、コピー中にデータを更新する場合は双方のボリュームのデータを更新する。ミラーリングはこのようにしてバックアップデータを得る技術である。

(第2の従来技術) バックアップデータを得る他の方法として、バックアップ処理中にデータが更新される場合、データの更新に先立ち、更新されるデータを他の領域に退避する方法が開示されている。バックアップ処理では更新されたデータに代えて退避データをバックアップする方法もある(例えば、特許文献2参照)。

(第3の従来技術) バックアップデータを得るさらに他の方法として、バックアップ対象のデータがあるボリュームの他に、そのボリューム(主ボリューム)と主副関係を持った仮想的な副ボリュームを用意し、バックアップ処理中に主ボリュームのデータが更新される場合、更新されるデータを副ボリュームに退避する方法がある(例えば、特許文献3参照)。更新されたデータに関するバックアップ処理は、第2の従来技術と同様である。

【0006】

第1の従来技術がコピー生成終了時点以降を特定時点としており、第2および第3の従来技術がデータ退避の開始時点を特定時点としているところが顕著な差である。

【0007】

【特許文献1】米国特許5,845,295号公報

【特許文献2】米国特許5,555,389号公報

【特許文献3】米国特許5,649,152号公報

【発明が解決しようとする課題】

バックアップデータのような特定時点のデータを凍結イメージのデータ又は単

に凍結イメージと呼ぶ。さらに凍結イメージをスナップショットと呼ぶことがある。凍結イメージは、バックアップデータとして保存しておく以外に、バッチ処理プログラム等のアプリケーションプログラムによって使用される。オンラインプログラム等の実行中に、これらのアプリケーションプログラムを実行することが必要な場合があり、特に24時間稼動を求められるコンピュータシステムでは、凍結イメージの生成はきわめて重要な課題である。

【0008】

このようなコンピュータシステムの利用環境において、第1の従来技術は既に述べたようにデータのコピーに要する時間が許容されない状況が発生している。また凍結イメージの取得および使用は、コピーの生成終了時点以降に限られるため、凍結イメージの生成を指示するコンピュータから見ると、特定時点の選択に制限があるという問題がある。

【0009】

一方、第2および第3の従来技術では、特定時点はホストコンピュータが凍結イメージ生成を指示し、データ退避を開始した時点となる。第2および第3の従来技術では、凍結イメージをバックアップデータとして保存する場合は、改めてコピー処理が必要となる。このコピー処理時には、ストレージ装置の制御負荷が集中し、オンラインプログラムの応答時間に影響を及ぼしかねない。

【0010】

そこで、本発明の目的は、ホストコンピュータの多様な使用環境においてもストレージあるいはホストに多大な負荷をかけずバックアップデータを取得するという利便性を持つストレージ装置およびその制御方法を提供することにある。

【0011】

本発明の他の目的は、ホストコンピュータの稼動状況への影響を少なくしたストレージ装置およびその制御方法を提供することにある。

本発明の更に他の目的は、本明細書および添付の図面の開示から明らかになる。

【0012】

【課題を解決するための手段】

上記課題を解決するために、本発明では、アクセスすべきデータの所在を示す論理スナップショット管理表を設け、論理的な凍結イメージ（スナップショット）の即時使用が可能な状態へ、单一のコマンドにより遷移することを可能とする。

【0013】

【発明の実施の形態】

以下、本発明の実施例を図面を用いて詳細に説明する。

【0014】

図1は、第1の実施例のシステム構成図である。図1において、1はストレージ装置、2はホストコンピュータ、3は管理用装置である。ストレージ装置1は制御装置11、プライマリボリューム13及びセカンダリボリューム14を含む。制御装置11は、図示を省略してあるがプロセッサ及びメモリを有し、内部で発生するイベントやホストコンピュータからの種々のコマンドやプライマリボリューム13及びセカンダリボリューム14へのアクセス（書き込み、読み出し）命令を、メモリ上の情報を参照しながら実行する。メモリ領域上の情報の一部として、本実施例では論理スナップショット管理表111がある。本実施例でフローチャートを用いて説明する処理は、いずれも制御装置11で実行される。

【0015】

ホストコンピュータ2は、図1ではストレージ装置1と直接接続しているが、ネットワークを介して接続しても良い。また、ホストコンピュータ2はファイルサーバ、パーソナルコンピュータ、ワークステーション等のコンピュータであり、複数台でも良い。管理用装置3は、ストレージ装置1の構成を定義したり、変更したりするための装置であり、コンピュータにその機能を持たせたものでも良い。本実施例では、後に詳述するホストコンピュータ2から発行される制御コマンドを管理用装置3から発行しても良い。ホストコンピュータ2からストレージ装置1に発行可能なコマンド、命令等は、管理用装置3からも発行できるようにしている。

【0016】

プライマリボリューム13及びセカンダリボリューム14の各々は、物理的な

ボリュームであっても論理的なボリュームであっても良い。いずれにしても本実施例では、プライマリボリューム13とセカンダリボリューム14とが対をなすものであれば良い。対は1対1の関係でなく、1:nの関係、すなわち一つのプライマリボリューム13に複数のセカンダリボリューム14を対応付けた構成であっても良い。さらに3つのボリュームを用意し、第1のボリュームを第1のプライマリボリューム、第2のボリュームを第1のプライマリボリュームに対応するセカンダリボリュームとすると共に第2のプライマリボリュームとし、第3のボリュームを第2のセカンダリボリュームとするカスケード関係にあっても良い。ここでは、説明を簡単にするためにセカンダリボリューム14が一つでプライマリボリューム13と対応付けられている場合について説明する。対応付けられた、プライマリボリューム13とセカンダリボリューム14との対をペアボリュームと呼ぶ。

【0017】

一例として、本実施例のプライマリボリューム13は、ホストコンピュータ2で実行されるオンラインプログラムのようなアプリケーションプログラムによってアクセスされる。セカンダリボリューム14は、ホストコンピュータ2で実行されるバックアッププログラムによってアクセスされる。プライマリボリューム13及びセカンダリボリューム14の各々の使い分けは、ホストコンピュータ2を用いる用途によって異なっても良いし、複数のホストコンピュータ2によって使い分けられても良い。

【0018】

なお、プライマリボリューム13及びセカンダリボリューム14のそれぞれは第1のボリューム、第2のボリューム、あるいは第1の記憶領域、第2の記憶領域と呼ぶことがある。

【0019】

図2は、本実施例におけるストレージ装置の制御を概観するための状態遷移図である。ストレージ装置1は、図2に示すSimplex状態210、Snap状態220およびDeleting状態230の3つの状態を有する。すなわち、複数のペアボリュームを持つストレージ装置1では、ペアボリュームごとに制御状態は異なる。ここ

では、説明を分かりやすくするために、ストレージ装置1には一つのペアボリュームがある場合を例に説明する。

【0020】

図2に矢印で示すように、幾つかのコマンド又はストレージ装置1の動作に応じて、これらの状態間に遷移が発生するように制御装置11により制御する。本実施例では、各状態はボリューム単位（ファイル単位等の他の単位であっても良い。）に制御するものとして、また説明を簡単にするために、ストレージ装置1には1対のプライマリボリューム13とセカンダリボリューム14とがあるものとして、説明する。図2に示す3つの状態を説明する。Simplex状態210は、プライマリボリューム13とセカンダリボリューム14とが独立した状態である。たとえば、ホストコンピュータ2のあるアプリケーションプログラム（たとえば、オンラインプログラム）による第1の記憶領域（プライマリボリューム13）へのアクセスと、他のアプリケーションプログラム（たとえば、バックアップ等のバッチ処理プログラム）による第2の記憶領域（セカンダリボリューム14）へのアクセスが何ら関係を持たずに実行可能な状態である。なお、厳密には、Simplex状態210ではこれらのボリュームはプライマリ、セカンダリという関係はないため、便宜的に第1、第2の記憶領域と呼称した。

【0021】

Snap状態220は、詳細は後述するが、プライマリボリューム13とセカンダリボリューム14とがペア状態をなし、制御上は主従関係にある状態である。

【0022】

Deleting状態230は、Snap状態220からSimplex状態210に遷移する際の過渡的な状態である。詳細は、Snap状態220の詳細とあわせて後述する。

【0023】

本実施例では、Snapコマンド240、Restoreコマンド260およびDeleteコマンド270の3種類の制御コマンドを、ホストコンピュータ2がストレージ装置1に向けて発行し、ストレージ装置1の制御装置11が実行するコマンドとして用意している。

【0024】

Snapコマンド240は、図2に示すように、第1、第2の記憶領域をSnap状態220にするためのコマンドである。Simplex状態210でSnapコマンド240が実行されると第1、第2の記憶領域は対となり、Snap状態220に遷移する。Snap状態220でSnapコマンド240が実行されるとペアボリュームはSnap状態220を保つ。

【0025】

Restoreコマンド260は、詳細は後述するが、Snap状態220でプライマリボリューム13とセカンダリボリューム14間でデータを操作するために発行されるコマンドであって、図2に示すようにペアボリュームの状態遷移を発生させるものではない。

【0026】

Deleteコマンド270は、図2ではSnap状態220からDeleting状態230への遷移を発生させるコマンドとして示してあるが、Deleting状態230が過渡的な状態であるので、Deleteコマンド270を発行するホストコンピュータ2から見れば、ペアボリュームをSnap状態220からSimplex状態210へ遷移させるコマンドである。

【0027】

なお、Deleting状態230を過渡的なものではなく、ストレージ装置の状態をホストコンピュータ2から分かるようにした上で、新たなコマンド等を用意し、ストレージ装置1の状態をDeleting状態230からSimplex状態210に遷移させることも可能である。

【0028】

図3にビットマップで表した論理スナップショット管理表111を示す。論理スナップショット管理表111は、本実施例で一例として示すビットマップに限らず、リスト構造や、物理アドレスを示すポインタ等であっても、管理情報または制御情報として、以下に説明するようなプライマリボリューム13とセカンダリボリューム14との対応付けの管理又は制御ができればよい。論理スナップショット管理表111で管理されるプライマリボリューム13とセカンダリボリューム14との対が前述のペアボリュームである。

【0029】

論理スナップショット管理表111は、Snapコマンド240の実行による、ストレージ装置1のSimplex状態210からSnap状態220への遷移に伴い作成される。論理スナップショット管理表111は、Deleteコマンド270の実行による、ストレージ装置1のSimplex状態210への遷移に伴い解放される。論理スナップショット管理表111を参照することにより、アクセスすべきデータの所在（物理アドレス）を知ることができる。

【0030】

プライマリボリュームBM(P)300がプライマリボリューム13のデータの保持状態を示し、セカンダリボリュームBM(S)301がセカンダリボリューム14のデータの保持状態を示す。図3に説明の便宜上付けた座標、たとえばプライマリボリュームBM(P)の(xp1, yp1)が示すプライマリボリューム13のデータとセカンダリボリュームBM(S)301の(xs1, ys1)が示すセカンダリボリューム14の記憶領域が対応する。プライマリボリュームBM(P)300、セカンダリボリュームBM(S)301夫々のビットにより表現されるデータの単位はブロック、ないしトラックなどであり、例えばデータの単位がトラックの場合、座標(xp1, yp1)の示すプライマリボリュームBM(P)、セカンダリボリュームBM(S)301はそれぞれプライマリボリューム13のシリンド番号xp1、ヘッド番号yp1で一意に定められるトラックのデータ状態、セカンダリボリューム14のシリンド番号xp1、ヘッド番号yp1で一意に定められるトラックのデータ状態を示す。

【0031】

図3は、例えば、Simplex状態210でSnapコマンド240が実行され、ストレージ装置1がSnap状態220に遷移した直後の論理スナップショット管理表111の内容を示す。このとき、プライマリボリューム13に記憶されている内容をスナップショット、あるいは凍結イメージ(static image)としてセカンダリボリュームに仮想的に取得する。すなわちセカンダリボリュームへ仮想的にスナップショットが存在するようにビットマップを用いて管理し、セカンダリボリュームへのアクセスにより、当該スナップショットにアクセスできるようにする。

【0032】

論理スナップショット管理表の一例であるビットマップにより、アクセス対象のデータの所在を知ることができる。セカンダリボリュームとプライマリボリュームが1:1に対応づけられた本実施例では、ビットマップの各ビットの意味は、「0」は対応する自ボリュームの領域にデータがあることを示し、「1」は対応する自ボリュームの領域にデータがないことを示す。

【0033】

したがって、図3に示す状態ではアクセスすべきデータはすべてプライマリボリューム13にある。この凍結状態にある時点のデータを、磁気テープ等の記憶媒体に格納することにより、その時点（Simplex状態210からSnap状態220に遷移した直後の時点）のバックアップがとれることになる。

【0034】

図4には以下に説明する、Snap状態220におけるアクセスすべきデータの所在の変化に伴う論理スナップショット管理表111の内容の一例を示す。

【0035】

論理スナップショット管理表111は、制御装置11内のメモリ上の設けられる制御情報として重要であり、メモリが揮発性メモリの場合は、不揮発性メモリに論理スナップショット管理表111のバックアップを取っておくことも必要である。

【0036】

以下の説明では、プライマリボリュームBM(P)の(x_pi, y_pj)の値(0又は1)とセカンダリボリュームBM(S)301の(x_si, y_sj)の値(0又は1)との関係をBM(P, S)と表記する。iおよびjは1, 2, 3, . . .であり、上限はボリューム各々のサイズで決まる値である。また以下の説明では、BM(P, S)の変化は、ホストコンピュータ2からのアクセスやコマンドによって生ずる。当該変化は、ビットマップを構成する各ビットによって管理されるデータの大きさごとに起こる。従って、ホストコンピュータ2からのアクセスのデータサイズとは必ずしも一致しない。

【0037】

Snap状態220においてビットマップは、ホストコンピュータ2からのプライマリボリューム13及びセカンダリボリューム14へのアクセス（書き込み、読み出し）命令や前述した制御コマンドによって変化する。以下、ビットマップの変化について説明する。

【0038】

図5は、ホストコンピュータ2からプライマリボリューム13またはセカンダリボリューム14へのアクセス（書き込み、読み出し）に応じたプライマリボリューム13及びセカンダリボリューム14のデータの保有状態の変化を、論理スナップショット管理表111のBM(P, S)の遷移として表したものである。図5の実行前の状態はRead（読み出し）又はWrite（書き込み）のアクセスを受ける前のBM(P, S)を表し、アクセス処理実行後の状態はRead（読み出し）又はWrite（書き込み）の処理完了後のBM(P, S)を表す。

【0039】

BM(P, S)は、(0, 0)、(0, 1)、(1, 0)および(1, 1)の4状態があるが、BM(P, S) = (1, 1)はプライマリボリューム13、セカンダリボリューム14のいずれにもアクセスすべきデータがない場合を示し、本実施例では存在し得ない。本実施例を具体化する上では、BM(P, S) = (1, 1)の状態で、ホストコンピュータ2からアクセスやコマンドを受けた場合は、それらを拒否する意味のエラー報告をホストコンピュータ2に送る。

【0040】

最初に、図5のBM(P, S) = (0, 1)の状態で、アクセスを受けた場合について説明する。BM(P, S) = (0, 1)はたとえば、Simplex状態210でSnapコマンド240が実行され、ペアボリュームがSnap状態220に遷移した直後に取りうる状態である。この状態のときにプライマリボリューム13へのRead処理が実行されても、すなわちホストコンピュータ2からのRead命令で指定されたアドレスからデータを読み出し、そのデータをホストコンピュータ2へ送っても、プライマリボリューム13及びセカンダリボリューム14のデータ保持状況のいずれも変化しないので、BM(P, S) = (0, 1)の状態が保たれる。

【0041】

図5のBM (P, S) = (0, 1) の状態における、プライマリボリューム13へのWrite処理を、図6を用いて説明する。ホストコンピュータ2からプライマリボリューム13へWrite命令が発行されると、制御装置11は論理スナップショット管理表111の、Write命令により特定されるプライマリボリューム13の領域に対応するBM (P, S) をチェックする（ステップ600）。BM (P, S) = (0, 1) であり、セカンダリボリューム14にアクセスすべきデータが存在しないので、プライマリボリューム13の領域にあるデータをセカンダリボリューム14の対応する領域へコピーする（ステップ610）。コピーを完了したら、論理スナップショット管理表111をBM (P, S) = (0, 1) からBM (P, S) = (0, 0) に更新する（ステップ615）。これらの処理の後に、Write命令で指示されるデータをプライマリボリューム13の所定の領域に書き込む（ステップ620）。

【0042】

上述のBM (P, S) = (0, 1) の状態における、プライマリボリューム13へのWrite処理において、プライマリボリューム13へデータ（新データ）を書き込む前に、その領域に格納されているデータ（旧データ）をセカンダリボリューム14の対応領域に退避（コピー）したことになる。したがって、Write処理完了後にプライマリボリューム13は新データに更新されているが、旧データはセカンダリボリューム14に格納されており、凍結イメージの時点のデータは保存されていることになる。

【0043】

ここで、論理スナップショット管理表111の各ビットが管理対象としているデータ領域のサイズ(たとえば、トラック単位)と書き込まれるデータ（新データ）のサイズとで前者の方が大きいか等しい場合は当該ビットが処理対象となり、書き込まれるデータの方が大きい場合は対応する複数ビットが処理対象になる。

【0044】

上記の図6の説明では、BM (P, S) = (0, 1) の状態を起点として説明したが、BM (P, S) = (1, 0) の状態を起点とする場合は、ステップ605及びステップ610におけるデータのコピー方向が逆になる。この点については後

述する。

【0045】

図5のBM (P, S) = (0, 1) の状態において、セカンダリボリューム14へのRead処理を、図7を用いて説明する。ホストコンピュータ2からセカンダリボリューム14へRead命令が発行されると、制御装置11は論理スナップショット管理表111の、Read命令により特定されるセカンダリボリューム14の領域に対応するBM (P, S) をチェックする（ステップ700）。BM (P, S) = (0, 1) であるので、セカンダリボリューム14にはアクセスすべきデータがなく（ステップ705）、アクセスすべきデータはプライマリボリューム13の対応する領域にあるので、その有意なデータをセカンダリボリューム14の当該領域にコピーする（ステップ710）。その後、論理スナップショット管理表111の (P, S) = (0, 1) を (P, S) = (0, 0) に更新する（ステップ715）。引き続いて、セカンダリボリューム14の当該領域にコピーしたデータを読み出し（ステップ720）、ホストコンピュータ2に送る。

【0046】

図7では、セカンダリボリューム14の当該領域にデータがない場合、セカンダリボリューム14にコピーしたデータを読み出し、ホストコンピュータに送るように説明したが、コピー動作の際に制御装置11にプライマリボリューム13からデータを読み出すので、その読み出したデータをホストコンピュータ2に送ることとしてもよい。これにより、ホストコンピュータ2から見ると、Read命令に対する応答時間が短くなる。Read命令に対してコピー動作を伴う同様の処理が以下でも説明されるが、以上と同様の処理により、応答時間を短くすることができる。

【0047】

セカンダリボリューム14の当該領域にアクセスすべきデータがある (P, S) = (0, 0) または (P, S) = (1, 0) の場合は、ステップ705においてステップ720に分岐し、当該データを読み出し、ホストコンピュータに送る。

【0048】

図7の処理において、セカンダリボリューム14にアクセスすべきデータがない場合に、対応するプライマリボリューム13のデータをセカンダリボリューム14にコピーすることを省略する処理であっても良い。この場合の処理を、図8を用いて説明する。図8のステップ800およびステップ805は、図7のステップ700およびステップ705と同じであるので、説明を省略する。ステップ805において、Read命令で特定される、セカンダリボリューム14の当該領域にデータがないならば($BM(P, S) = (0, 1)$)、プライマリボリューム13の対応する領域のデータを読み出し(ステップ810)、ホストコンピュータ2に送る。一方、セカンダリボリューム14にアクセスすべきデータがあるならば($BM(P, S) = (0, 0)$ 又は $(P, S) = (1, 0)$)、セカンダリボリューム14の当該領域のデータを読み出し(ステップ815)、ホストコンピュータに送る。

【0049】

図8の処理によっても、バックアップをとるという観点では、図7の処理と同じになる。図8の処理では、セカンダリボリューム14にコピーされているデータは、プライマリボリューム13のうち、Snapコマンド受領後に更新された領域の旧データと、Snapコマンド受領後に更新されたセカンダリボリューム14のデータのみであるので、セカンダリボリューム14の容量は少なくて済む。言い換えるれば、凍結イメージがプライマリボリューム13にあるデータについては、対応する領域をセカンダリボリューム14に確保しなくて済む。具体的には、セカンダリボリュームの容量を、用途等を考慮した統計的予測手法を用いて決定し、容量を削減することが可能となる。また、セカンダリボリューム14へのコピーが必要な都度、領域を確保することでも小容量のセカンダリボリューム14を実現できる。

【0050】

図7及び図8の処理は、ストレージ装置1の役割、すなわちホストコンピュータ2の用途や機能により使い分けることが望ましい。そこで、本実施例では、制御装置11に図7及び図8双方の処理機能を持たせ、ユーザによってホストコンピュータ2又は管理用装置3から選択できるようにしている。

【0051】

本実施例はSnap状態においても、ホストコンピュータ2からセカンダリボリューム14へのWrite命令の発行を許す。

図5のBM(P, S) = (0, 1)の状態において、セカンダリボリューム14へのWrite処理を、図9を用いて説明する。ホストコンピュータ2からセカンダリボリューム14へWrite命令が発行されると、制御装置11は論理スナップショット管理表111の、Write命令により特定されるセカンダリボリューム14の領域のBM(P, S)をチェックする(ステップ900)。BM(P, S) = (0, 1)であるので、セカンダリボリューム14の当該領域にはアクセスすべきデータがなく(ステップ905)、プライマリボリューム13の対応する領域にアクセスすべきデータがあるので、当該データをセカンダリボリューム14の当該領域にコピーする(ステップ910)。その後、論理スナップショット管理表111の(P, S) = (0, 1)を(P, S) = (0, 0)に更新する(ステップ915)。引き続いて、セカンダリボリューム14の当該領域に、Write命令で指示されるデータを書き込む(ステップ920)。

【0052】

以上の処理において、コピーしたデータの上にWrite命令で指示されるデータを書き込むことは、前述したように論理スナップショット管理表111の各ビットが管理対象としているデータのサイズよりもWrite命令で指示されるデータのサイズが小さく、結果としてコピーしたデータの一部を更新する場合に対応するためである。論理スナップショット管理表111の複数ビットに渡るデータのサイズが指示された場合は、各ビット対応のサイズにデータを分割して実行する。

【0053】

なお、Write命令で指示されるデータのサイズと論理スナップショット管理表111の各ビットが管理対象としているデータのサイズとが等しい場合は、ステップ910のコピー処理を省略できる。

【0054】

セカンダリボリューム14にもアクセスすべきデータがある(P, S) = (0, 0)の場合は、ステップ905においてステップ920に分岐し、当該データの

上に、Write命令で指示されるデータを書き込む。

【0055】

図9の処理を許すことにより、凍結イメージのデータを用いた、ホストコンピュータ2でセカンダリボリューム14へのWrite命令の発行を含むアプリケーションプログラムのテストの実行や種々のプログラムの実行を可能にする。プログラムの実行の例としては、金融機関の営業終了時間の凍結イメージを基にその日の日計処理（1日の種々の営業統計処理）がある。このような処理ではボリュームのRead処理が主体であるが、ソート処理のようにWrite処理が発生する場合がある。

【0056】

上記の図9の説明では、 $BM(P, S) = (0, 1)$ の状態を起点として説明したが、 $BM(P, S) = (1, 0)$ の状態を起点とする場合は、ステップ905及びステップ910におけるデータのコピー方向が逆になる。この点については後述する。

【0057】

図5の $BM(P, S) = (0, 0)$ の状態における、プライマリボリューム13又はセカンダリボリューム14へのアクセスに伴う処理を説明する。 $BM(P, S) = (0, 0)$ の状態は、図6～図9を用いて説明したように、Snapコマンド240が実行されSimplex状態210から遷移したSnap状態220で、たとえばプライマリボリューム13に対してWrite処理が実行された状態である。

【0058】

図5の $BM(P, S) = (0, 0)$ の状態において、プライマリボリューム13及びセカンダリボリューム14の各々へのRead命令とWrite命令とのいずれかの命令の発行パターンは4通りある。 $BM(P, S) = (0, 0)$ はプライマリボリューム13及びセカンダリボリューム14のいずれのデータも有意であるので、データが維持されるRead処理だけでなくWrite処理に於いても、データが更新され、更新されたデータが有意であるので、論理スナップショット管理表111は $BM(P, S) = (0, 0)$ のまま維持される。

【0059】

以上の論理スナップショット管理表111が変化しない処理のうち、プライマリボリューム13からのRead処理は後述するが、プライマリボリューム13へのWrite処理では、図6のステップ605において、Write対象の領域のデータは既にセカンダリボリューム14へコピー済みであるからステップ625へ分岐し、状態遷移を管理する論理スナップショット管理表111の更新は不要であるので、ステップ625からステップ620へ分岐する。

【0060】

セカンダリボリューム14からのRead処理では、図7のステップ705でセカンダリボリューム14はアクセス（Read処理）対象のデータを保持しており、ステップ720へ分岐し、論理スナップショット管理表111の更新を伴わない。図7の他の方法を示した図8では、論理スナップショット管理表111の更新の処理を含まないのは図面から明らかである。

【0061】

セカンダリボリューム14へのWrite処理では、図9のステップ905において、Write対象の領域のデータは既にセカンダリボリューム14へコピー済みであるからステップ920へ分岐し、論理スナップショット管理表111の更新を伴わない。

【0062】

次に、図5のBM（P, S）=（1, 0）の状態における、プライマリボリューム13又はセカンダリボリューム14へのアクセスに伴う処理を説明する。BM（P, S）=（1, 0）は、BM（P, S）=（0, 0）の状態において後述するRestoreコマンドが発行されたときの状態である。BM（P, S）=（1, 0）は、プライマリボリューム13にアクセスすべきデータがなく、セカンダリボリューム14にアクセスすべきデータがある状態を示す。

【0063】

図5のBM（P, S）=（1, 0）の状態において、プライマリボリューム13のRead処理を、図10を用いて説明する。ホストコンピュータ2からプライマリボリューム13へRead命令が発行されると、制御装置11は論理スナップショット管理表111の、Read命令により特定されるプライマリボリューム13の領域

のBM (P, S) をチェックする（ステップ1000）。BM (P, S) = (1, 0) であるので、プライマリボリューム13にはアクセスすべきデータはなく（ステップ1005）、セカンダリボリューム14の対応する領域にアクセスすべきデータがあるので、当該データをプライマリボリューム13の当該領域にコピーする（ステップ1010）。その後、論理スナップショット管理表111の (P, S) = (1, 0) を (P, S) = (0, 0) に更新する（ステップ1015）。引き続いて、Read命令で特定される、プライマリボリューム13の領域からデータを読み出し（ステップ1020）、ホストコンピュータに送る。

【0064】

前述のBM (P, S) = (0, 1) 及びBM (P, S) = (0, 0) におけるプライマリボリューム13のデータのRead処理について前者は図面を用いずに説明し、後者は説明を保留していたが、これらの処理は、図10においてステップ1005からステップ1020へ分岐する処理になることは容易に理解される。

【0065】

図5のBM (P, S) = (1, 0) の状態における、プライマリボリューム13のWrite処理を図6を用いて説明する。図6のデータのコピー方向は、前述したように、BM (P, S) = (0, 1) の状態を起点とした場合と逆になる。BM (P, S) = (1, 0) の状態において、ホストコンピュータ2からプライマリボリューム13へWrite命令が発行されると、制御装置11は論理スナップショット管理表111の、Write命令により特定されるプライマリボリューム13の領域のBM (P, S) をチェックする（ステップ600）。BM (P, S) = (1, 0) であり、プライマリボリューム13の対応する領域のデータは無意であるので、セカンダリボリューム14の領域にあるデータをプライマリボリューム13の対応する領域へコピーする（ステップ610）。コピーを完了したら、論理スナップショット管理表111をBM (P, S) = (1, 0) からBM (P, S) = (0, 0) に更新する（ステップ615）。この後に、Write命令で指示されるデータをプライマリボリューム13の所定の領域に書き込む。

【0066】

図5のBM (P, S) = (1, 0) の状態における、セカンダリボリューム14

からのRead処理では、図7のステップ705でセカンダリボリューム14はアクセス（Read処理）対象のデータを保持しており、ステップ720へ分岐し、ホストコンピュータ2からのRead命令で指定されたアドレスからデータを読み出し、そのデータをホストコンピュータ2へ送る（ステップ720）。図7の他の方法として示した図8では、ステップ805で分岐し、Read命令で指定されたアドレスからデータを読み出し、そのデータをホストコンピュータ2へ送る（ステップ815）。

【0067】

図5のBM（P, S）=（1, 0）の状態における、セカンダリボリューム14へのWrite処理を図9を用いて説明する。図9のデータのコピー方向は、前述したように、BM（P, S）=（0, 1）の状態を起点とした場合と逆になる。BM（P, S）=（1, 0）の状態において、ホストコンピュータ2からセカンダリボリューム14へWrite命令が発行されると、制御装置11は論理スナップショット管理表111の、Write命令により特定されるセカンダリボリューム14の領域のBM（P, S）をチェックする（ステップ900）。BM（P, S）=（1, 0）であり、プライマリボリューム13の対応する領域にはアクセスすべきデータがないので、セカンダリボリューム14の領域にあるデータをプライマリボリューム13の対応する領域へコピーする（ステップ910）。コピーを完了したら、論理スナップショット管理表111をBM（P, S）=（1, 0）からBM（P, S）=（0, 0）に更新する（ステップ915）。この後に、Write命令で指示されるデータをセカンダリボリューム14の所定の領域に書き込む。

【0068】

本実施例では、Snapコマンド240、Restoreコマンド260およびDeleteコマンド270の3種類の制御コマンドを、ホストコンピュータ2が発行するコマンドとして用意していることを既に述べた。Deleteコマンド270は、既に説明したように、Snap状態220からDeleting状態230への遷移を発生させるコマンドであるので、アクセスすべきデータの所在の変化を生起させない。ペアボリュームがSnap状態220にあるときに、Snapコマンド240またはRestoreコマンド260がホストコンピュータ2から発行されると、Snap状態220が維持さ

れるが、アクセスすべきデータの所在の変更が発生する。

【0069】

Snapコマンド240は、プライマリボリューム13にアクセスすべきデータがあれば、セカンダリボリューム14の対応するデータを無意にし、セカンダリボリューム14のアクセスすべきデータの所在がプライマリボリューム13のアクセスすべきデータの所在と等しくなるように論理スナップショット管理表111を更新する。セカンダリボリューム14にのみアクセスすべきデータがあれば、その状態を維持する。**Restore**コマンド260は、**Snap**状態220にあるときに、プライマリボリューム13内のデータを何らかの理由で使用できなくなった場合、たとえばホストコンピュータ2が正しいデータに誤ったデータを上書きしてしまったような場合、ホストコンピュータ2からのリカバリーのために用意したコマンドである。

【0070】

Restoreコマンド260は、セカンダリボリューム14にアクセスすべきデータがあれば、プライマリボリューム13の対応するデータを無意にし、プライマリボリューム13のアクセスすべきデータの所在がセカンダリボリューム14のアクセスすべきデータの所在と等しくなるように論理スナップショット管理表111を更新する。プライマリボリューム13にのみアクセスすべきデータがあれば、その状態を維持するためのコマンドである。このようなアクセスすべきデータの所在の変化について説明した後に、**Delete**コマンド270に対応する処理を説明する。

【0071】

図11は、ホストコンピュータ2からプライマリボリューム13またはセカンダリボリューム14への**Snap**コマンド240または**Restore**コマンド260の発行に応じたプライマリボリューム13及びセカンダリボリューム14のデータの保有状態の変化を、論理スナップショット管理表111のBM(P, S)の変化として表したものである。図11の実行前の状態は**Snap**コマンド240または**Restore**コマンド260を受ける前のBM(P, S)を表し、アクセス処理実行後の状態は**Snap**コマンド240または**Restore**コマンド260を受け、所定の処理完了

後のBM (P, S) を表す。

【0072】

実行前のBM (P, S) は、(0, 0)、(0, 1) および(1, 0) の3状態を取り得る。BM (P, S) = (1, 1) はプライマリボリューム13及びセカンダリボリューム14のいずれもがアクセスすべきデータを保有しない場合を示し、本実施例では存在しない。

【0073】

以下の説明で明らかになるように、Snapコマンド240またはRestoreコマンド260が発行されると、制御装置11はペアボリューム内の全データについて処理を実行する。ペアボリューム内のデータに関して、実行前のBM (P, S) が異なれば、処理が異なるので、実行前のBM (P, S) の状態ごとに、以下説明する。

【0074】

図11のBM (P, S) = (0, 1) の状態で、Snapコマンド240を受けた場合について図12を用いて説明する。BM (P, S) = (0, 1) の状態はたとえば、図3で説明した、Simplex状態210でSnapコマンド240が実行され、ペアボリュームがSnap状態220に遷移した直後に現れる状態である。論理スナップショット管理表111をチェックすると（ステップ1200）、セカンダリボリューム14にはアクセスすべきデータが存在しないので（ステップ1205）、ステップ1220に分岐する。

【0075】

図12では、論理スナップショット管理表111で管理するすべてのデータをチェックするための処理ループ（ステップ1220からステップ1200への分岐）を制御するための初期設定を省略してあるが、ステップ1220で全データに関する処理の終了を判断し、終了していなければステップ1200へ分岐する。以上の処理により、BM (P, S) = (0, 1) の状態で、Snapコマンド240を受けると、ペアボリュームはBM (P, S) = (0, 1) の状態を維持する。

【0076】

図11のBM (P, S) = (0, 1) の状態で、Restoreコマンド260を受け

た場合について図13を用いて説明する。論理スナップショット管理表111をチェックすると（ステップ1300）、プライマリボリューム13にアクセスすべきデータが存在し（ステップ1305）、セカンダリボリューム14の対応する領域にはアクセスすべきデータが存在しないので（ステップ1310）、ステップ1320に分岐する。図13でも、論理スナップショット管理表111で管理するすべてのデータをチェックするための処理ループ（ステップ1320からステップ1300への分岐）を制御するための初期設定を省略してあるが、ステップ1320で全データに関する処理の終了を判断し、終了していないければステップ1300へ分岐する。以上の処理により、 $BM(P, S) = (0, 1)$ の状態で、Restoreコマンド260を受けると、ストレージ装置1は $BM(P, S) = (0, 1)$ の状態を維持する。

【0077】

図11の $BM(P, S) = (0, 0)$ の状態で、Snapコマンド240を受けた場合について図12を用いて、既に説明したことと異なる部分に関して主に説明する。セカンダリボリューム14にアクセスすべきデータがあり（ステップ1205）、プライマリボリューム13の対応するデータも有意であるので（ステップ1210）、論理スナップショット管理表111の $BM(P, S) = (0, 0)$ を $BM(P, S) = (0, 1)$ に更新する（ステップ1215）。以上の処理を全データに関して繰り返す（ステップ1220）。

【0078】

図11の $BM(P, S) = (0, 0)$ の状態で、Restoreコマンド260を受けた場合について図13を用いて、既に説明したことと異なる部分に関して主に説明する。プライマリボリューム13のデータはコマンド受領まで有意であり（ステップ1305）、セカンダリボリューム14の対応する領域にもアクセスすべきデータが存在するので（ステップ1310）、論理スナップショット管理表111の $BM(P, S) = (0, 0)$ を $BM(P, S) = (1, 0)$ に更新する（ステップ1315）。以上の処理を全データに関して繰り返す（ステップ1320）。

【0079】

図11のBM (P, S) = (1, 0) の状態で、Snapコマンド240を受けた場合について図12を用いて説明する。セカンダリボリューム14にアクセスすべきデータがあり（ステップ1205）、プライマリボリューム13の対応する領域にはアクセスすべきデータがないので（ステップ1210）、ステップ1220へ分岐する。BM (P, S) = (1, 0) の状態は維持される。以上の処理を全データに関して繰り返す（ステップ1220）。

【0080】

図11のBM (P, S) = (1, 0) の状態で、Restoreコマンド260を受けた場合について図13を用いて説明する。プライマリボリューム13にはアクセスすべきデータが存在しないので（ステップ1305）、ステップ1220へ分岐する。BM (P, S) = (1, 0) の状態は維持される。以上の処理を全データに関して繰り返す（ステップ1320）。

【0081】

Restoreコマンド260を実行後に、BM (P, S) = (1, 0) である領域については、ホストコンピュータ2からのアクセスとは非同期に、セカンダリボリューム14の当該領域に保持されているデータを、プライマリボリューム13の対応する領域にコピーしてもよい。この動作を図14を用いて説明する。

【0082】

Restoreコマンド260を実行した結果、すべての領域はBM (P, S) = (1, 0) もしくはBM (P, S) = (0, 1) のいずれかの状態になっている。論理スナップショット管理表111をチェックし（ステップ1400）、プライマリボリューム13にアクセスすべきデータが存在するならば（BM (P, S) = (0, 1)）（ステップ1405）、ステップ1420へ分岐する。

【0083】

一方、プライマリボリューム13にアクセスすべきデータがないならば（BM (P, S) = (1, 0)）、セカンダリボリューム14の対応する領域には有意なデータが必ず保持されているので、当該データを当該領域にコピーする（ステップ1410）。その後、論理スナップショット管理表111をBM (P, S) = (0, 0) にする（ステップ1415）。以上の処理を全データに関して繰り返す

(ステップ1420)。

【0084】

このコピー動作により、ホストコンピュータ2からのアクセスの有無によらず、プライマリボリューム13のすべてのデータが有意なものとなる。

【0085】

以上より、直前の制御コマンドがSnapコマンド240であれば、Snapコマンド240実行時点のプライマリボリューム13のデータがセカンダリボリューム14へのアクセスに対して保証され、直前の制御コマンドがRestoreコマンド260であれば、Restoreコマンド260実行時点のセカンダリボリューム14のデータがプライマリボリューム13へのアクセスに対して保証される。また、Snapコマンド240実行後に、更新されたセカンダリボリューム14のデータはSnapコマンド240の実行によりセカンダリボリューム14に保存されたデータより優先され、Restoreコマンド260実行後に、更新されたプライマリボリューム13のデータはRestoreコマンド260実行時にプライマリボリューム13にコピーされたデータより優先される。

【0086】

すなわち、かかる場合にはコマンド実行時点のデータを更新データで上書きし、コマンド実行時点のデータは保証されない。

【0087】

次にDeleteコマンド270について説明する。既に説明したように、Deleteコマンド270は、Snap状態220から、過渡的なDeleting状態230を経由して、Simplex状態210へ遷移させるコマンドである。Deleteコマンドの受領によっては論理スナップショット管理票111は変化しない。この状態を図15に示す。図16に制御装置が実行するDelete動作を示す。

【0088】

Deleteコマンド270を受けると、プライマリボリューム13にアクセスすべきデータが無いならば(ステップ1600)、ステップ1615へ分岐し、アクセスすべきデータがあるならば、セカンダリボリューム14の対応する領域にアクセスすべきデータがあるかをチェックする(ステップ1605)。アクセスすべき

データがあるならばステップ1620へ分岐する。セカンダリボリューム14の対応する領域にアクセスすべきデータがないならば、プライマリボリューム13からセカンダリボリューム14へ当該データをコピーする(ステップ1610)。ステップ1615では、セカンダリボリューム14の当該データをプライマリボリューム13の対応する領域へコピーする。ステップ1620においては、プライマリボリューム13及びセカンダリボリューム14のいずれもがアクセスすべきデータを保有しているので、論理スナップショット管理表111をBM(P, S) = (0, 0)とする。以上の処理を全データに関して繰り返す(ステップ1625)。

【0089】

なお、Simplex状態210で論理スナップショット管理表111による管理が不要である場合は、処理済みアドレスをポインタ等により管理することにより、図16のステップ1620を省略できる。

【0090】

Deleting状態230を過渡的なものではなく、ホストコンピュータ2から分かる状態にして、新たなコマンド等を用意し、ストレージ装置1の状態をDeleting状態230からSimplex状態210に遷移させることも可能であることを前述した。ここでは、Snap状態220からミクロな状態はそのままにして、マクロな状態をDeleting状態230にする。Deleting状態230であることをメモリに記憶しておけば、Snap状態220との状態の差異を制御装置11は認識できる。また、必要に応じてホストコンピュータ2や管理用装置3にDeleting状態230であることを報告する。Deleting状態230において、ホストコンピュータ2からのプライマリボリューム13またはセカンダリボリューム14へのアクセスに応じて、Simplex状態210に遷移させる方法について以下に説明する。

【0091】

この場合の、論理スナップショット管理表111の変化は既に説明した図5に従う。Simplex状態210に遷移する直前には、論理スナップショット管理表111はBM(P, S) = (0, 0)になっている必要があるが、図5に従えば、BM(P, S) = (0, 1)の状態でプライマリボリューム13にRead命令が発行さ

れた場合とBM(P, S) = (1, 0)の状態でセカンダリボリューム14にRead命令が発行された場合では、BM(P, S) = (0, 0)への状態の変化は起こらない。このようなデータが存在する場合は、ストレージ装置1はDeleting状態230を維持し、BM(P, S) = (0, 0)への変化を生起させるアクセスの実行に伴って、Simplex状態210に遷移する。

【0092】

以上の第1実施例により、データの凍結イメージ(スナップショット)へのアクセスが、Snapコマンド240を実行後に直ちに可能になる。またSnapコマンド240を実行後は、各データの凍結イメージが論理的なスナップショット(セカンダリボリュームがデータを保持していない状態)又は物理的なスナップショット(セカンダリボリュームがデータを保持している状態)で維持されているので、ホストコンピュータからのプライマリボリュームへのアクセスは、セカンダリボリュームへのアクセスによって制約を受けることは無い。

【0093】

さらにRestoreコマンド260の実行により、プライマリボリュームのデータに誤りが生じても、セカンダリボリュームのデータを用いて復元することができる。Restoreコマンド260はSnapコマンド240と対称的な制御をするために用意されているので、ストレージ装置1の制御に関わらず、ホストコンピュータ2にはプライマリボリューム13とセカンダリボリューム14との役割を逆にする自由度を与えていている。

【0094】

ホストコンピュータ2からのDeleteコマンド270の発行に応じて、プライマリボリューム13とセカンダリボリューム14の双方がデータを保有している状態にした後に、各々のボリュームを独立して使用できるSimplex状態210へ遷移するので、各ボリュームの利用性が高まる。

【0095】

データのバックアップを保存する場合、用途により保存の仕方が変わる。毎月1回のバックアップや1年を超えるような保存期間のバックアップの場合は、磁気テープ等にバックアップデータを保存しておく必要がある。一方、データの誤

り回復のための一時的なバックアップは、10分毎や1時間毎に実行される。この場合、磁気テープ等の記憶媒体を新たに用意する必要はなく、本実施例のSimplex状態210に戻ったセカンダリボリューム14をバックアップ用として使用できる。ある時点のバックアップのあるセカンダリボリューム14に、本実施例により次のバックアップデータを上書きすることにより、バックアップ用の記憶容量を増加させないで済む。バックアップデータの上書き中に、セカンダリボリューム14に障害が発生すると回復不能になる場合があるので、セカンダリボリュームを2つ用意し、交互に使用すれば、回復不能になる状態を避けることができる。

【0096】

本発明の第2実施例を説明する。第1実施例では、Snap状態220からSimplex状態210へ遷移のために、Deleteコマンド270を実行する。ところが、Simplex状態210へ遷移した後は、セカンダリボリューム14のデータを使用しない用途も多々ある。典型的にはSnap状態220でデータのバックアップを完了したら、以後はセカンダリボリューム14のデータを使用しないというものである。このような用途のためには、Deleteコマンド270の実行に伴うデータのコピーは不要である。不要なデータのコピーに時間を要しているという問題もある。

【0097】

そこで、図17に示すように、Deleteコマンド270とは異なる、データコピーをせずに短時間でプライマリボリューム13とセカンダリボリューム14とのペア関係を解消するDivideコマンド1300を用意する。ストレージ装置がSimplex状態210へ遷移したことをホストコンピュータ2が認識できれば良く、制御装置11によるDivideコマンド1300の処理としては、論理スナップショット管理表111を解放し、ホストコンピュータ2へ報告すればよい。

【0098】

Divideコマンド1300を用意した本実施例により、ストレージ装置をSnap状態220からSimplex状態210へ短時間で遷移させることができる。

【0099】

Snap状態220におけるセカンダリボリューム14へのアクセスは、データの

バックアップを取るときのようにRead命令に限られる場合は、本実施例の短時間でペア関係を解消することは極めて有用である。

【0100】

Deleteコマンド270を用いずに、Divideコマンド1300を使用する場合は、セカンダリボリューム14の容量がプライマリボリューム13の容量に比べて少なくてすむ。セカンダリボリューム14の容量は、その用途等を考慮した統計的手法を用いることにより、適切に定めることができる。

【0101】

第1実施例及び第2実施例を合わせて考慮すると、Snap状態220からSimple x状態210への遷移には、Deleteコマンド270及びDivideコマンド1300のいずれか一方を用いる制御と、それらの双方を用いる制御の3通りの制御モードがある。例えばDivideコマンド1300のみを用いた例を図18に示す。

【0102】

これらの制御モードの選択は、幾つかの例で示してきたようにホストコンピュータ2の役割や機能に依存する。そこで、ホストコンピュータ2の役割や機能が予め定まっている場合は、それに応じた特定の制御モードをストレージ装置1が持てば良いが、時間経過と共に必要な制御モードが変化することが多い。そこで、制御モードはホストコンピュータ2から選択できるようにする。ホストコンピュータ2や管理用装置3からの制御モードの選択の方法は、制御モードを指定するための新たなコマンドを設けても良いが、Snapコマンド240のパラメータとして付加する。制御装置11はそのパラメータをメモリに記憶、参照しながら制御を実行する。

【0103】

このように制御モードの選択をホストコンピュータ2を用いるユーザに委ねることにより、ユーザはホストコンピュータ2の用途や機能に応じた好適なストレージ装置を利用できる。

【0104】

【発明の効果】

以上説明したように本発明のストレージ装置およびその管理方法によれば、ホ

ストコンピュータの使用環境に応じた最適な凍結イメージの使用および取得が可能となり、ストレージ装置の利便性を向上することができる。

【図面の簡単な説明】

【図1】

本発明のストレージ装置の構成図である。

【図2】

本発明のストレージ装置が取り得る状態間の状態遷移図である。

【図3】

本発明の、論理スナップショット管理表の一例を示す図である。

【図4】

本発明の、論理スナップショット管理表の一例を示す図である。

【図5】

本発明の論理スナップショット管理表の、アクセス処理に伴う変化を示す図である。

【図6】

本発明の、プライマリボリューム更新時のフローチャートを示す図である。

【図7】

本発明の、セカンダリボリューム読み出し時のフローチャートを示す図である。

【図8】

本発明の、セカンダリボリューム読み出し時のフローチャートを示す図である。

【図9】

本発明の、セカンダリボリューム更新時のフローチャートを示す図である。

【図10】

本発明の、プライマリボリューム読み出し時のフローチャートを示す図である。

【図11】

本発明の論理スナップショット管理表の、制御コマンド受領に伴う変化を示す図である。

【図12】

本発明の、Snapコマンド実行のフローチャートを示す図である。

【図13】

本発明の、Restoreコマンド実行のフローチャートを示す図である。

【図14】

本発明の、Restore動作のフローチャートを示す図である。

【図15】

本発明の、Deleteコマンド実行前後の論理スナップショット管理表の状態を示す図である。

【図16】

本発明の、Delete動作のフローチャートを示す図である。

【図17】

本発明のストレージ装置が取り得る状態間の状態遷移図である。

【図18】

本発明のストレージ装置が取り得る状態間の状態遷移図である。

【符号の説明】

1 ……ストレージ装置

2 ……ホストコンピュータ

3 ……管理用装置

1 1 ……制御装置

1 3 ……プライマリボリューム

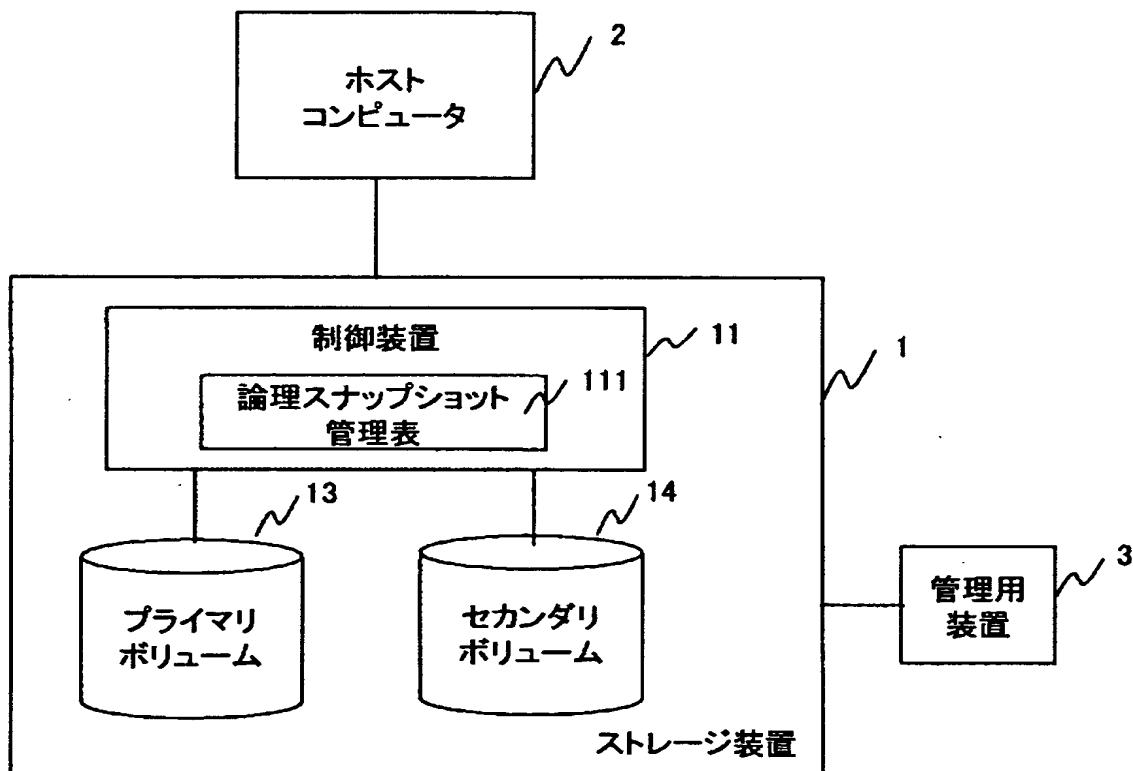
1 4 ……セカンダリボリューム

1 1 1 ……論理スナップショット管理表

【書類名】 図面

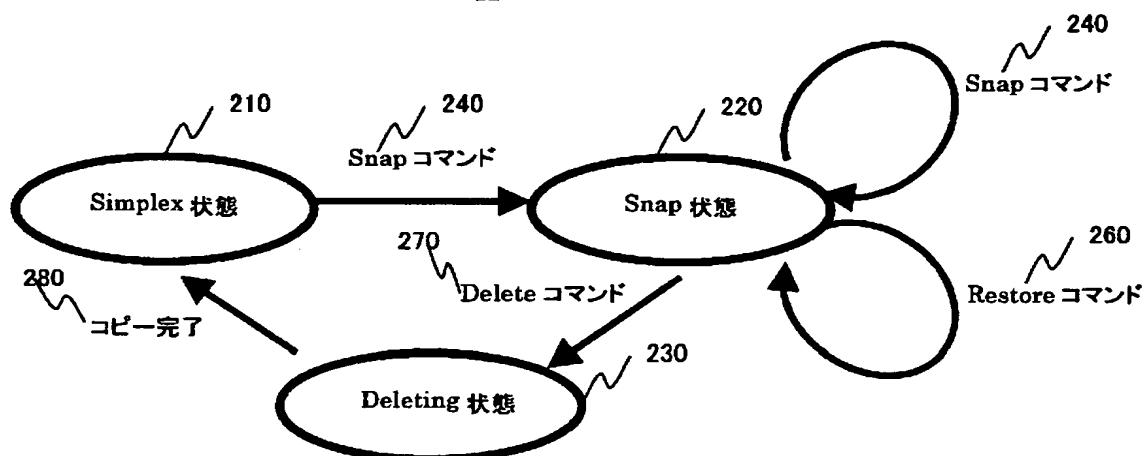
【図1】

図1



【図2】

図2



【図3】

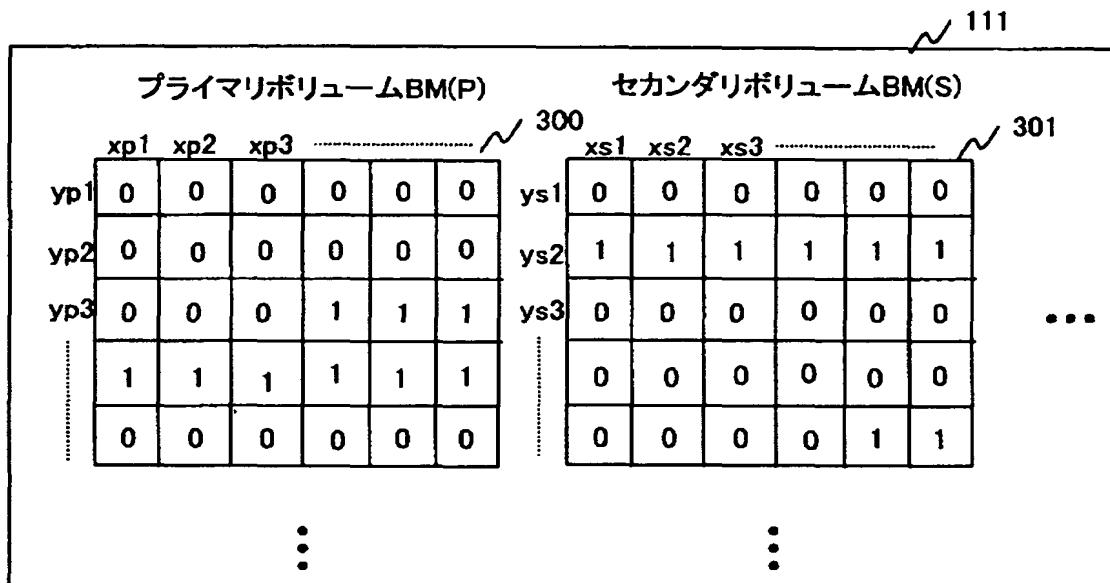
図3

111

プライマリボリュームBM(P)						セカンダリボリュームBM(S)					
	xp1	xp2	xp3		xs1	xs2	xs3		
yp1	0	0	0	0	0	0	1	1	1	1	1
yp2	0	0	0	0	0	0	1	1	1	1	1
yp3	0	0	0	0	0	0	1	1	1	1	1
.....	0	0	0	0	0	0	1	1	1	1	1
.....	0	0	0	0	0	0	1	1	1	1	1
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図4】

図4



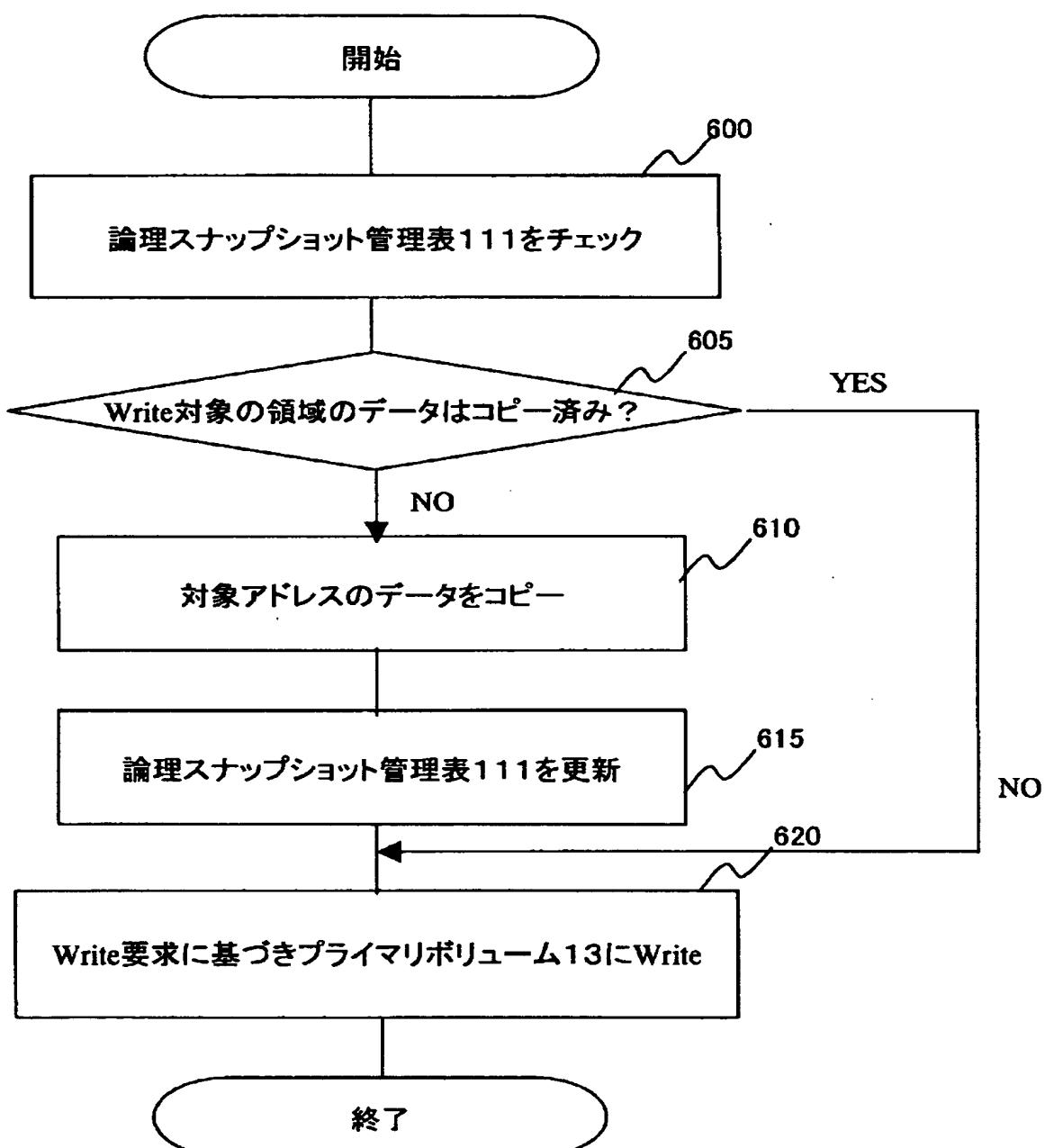
【図5】

図5

実行前の状態 BM (P,S)	プライマリボリュームへのアクセス		セカンダリボリュームへのアクセス	
	Read	Write	Read	Write
	(0, 0)	(0, 0)	(0, 0)	(0, 0)
(0, 1)	(0, 1)	(0, 0)	(0, 0)	(0, 0)
(1, 0)	(0, 0)	(0, 0)	(1, 0)	(0, 0)
(1, 1)	N/A			

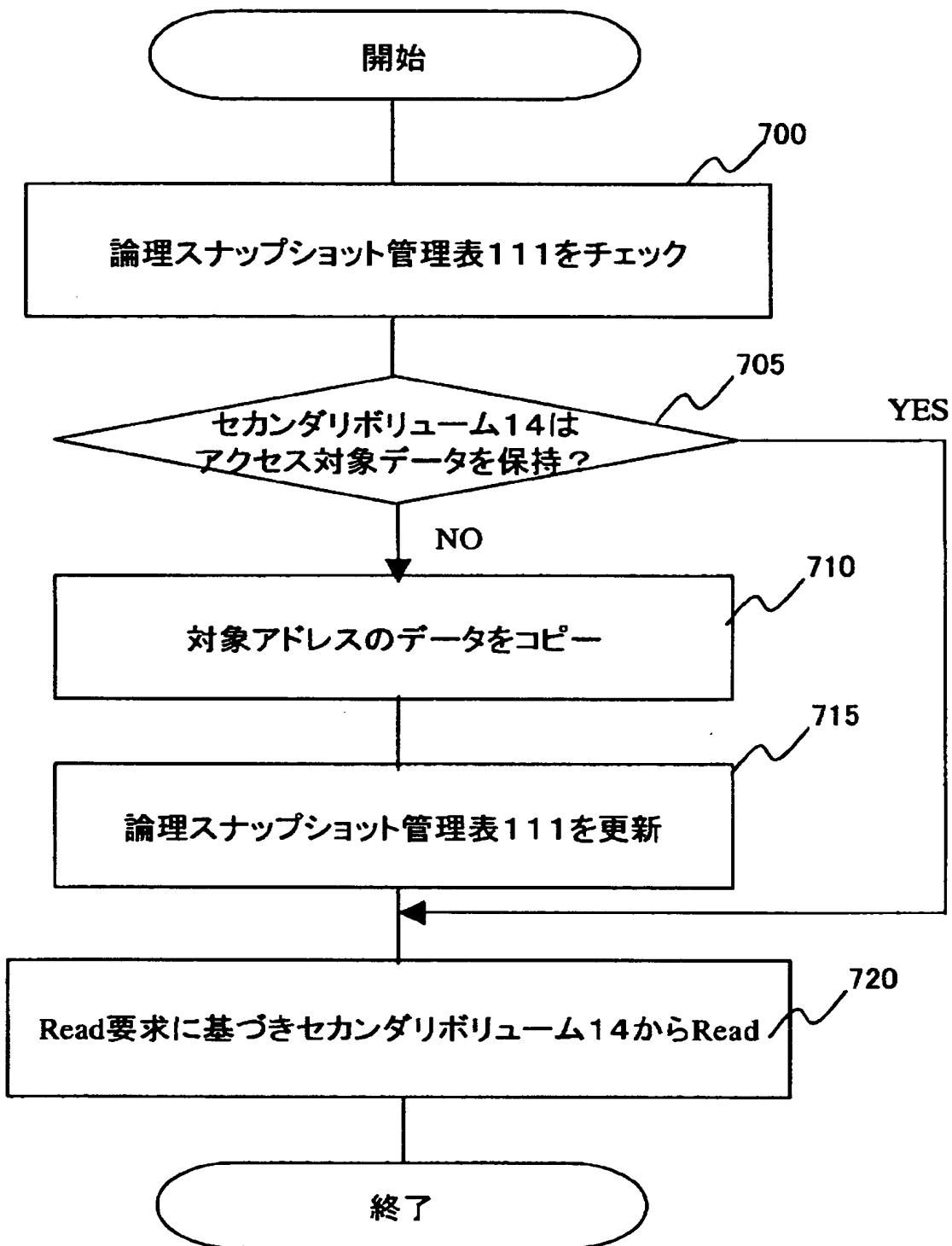
【図6】

図6



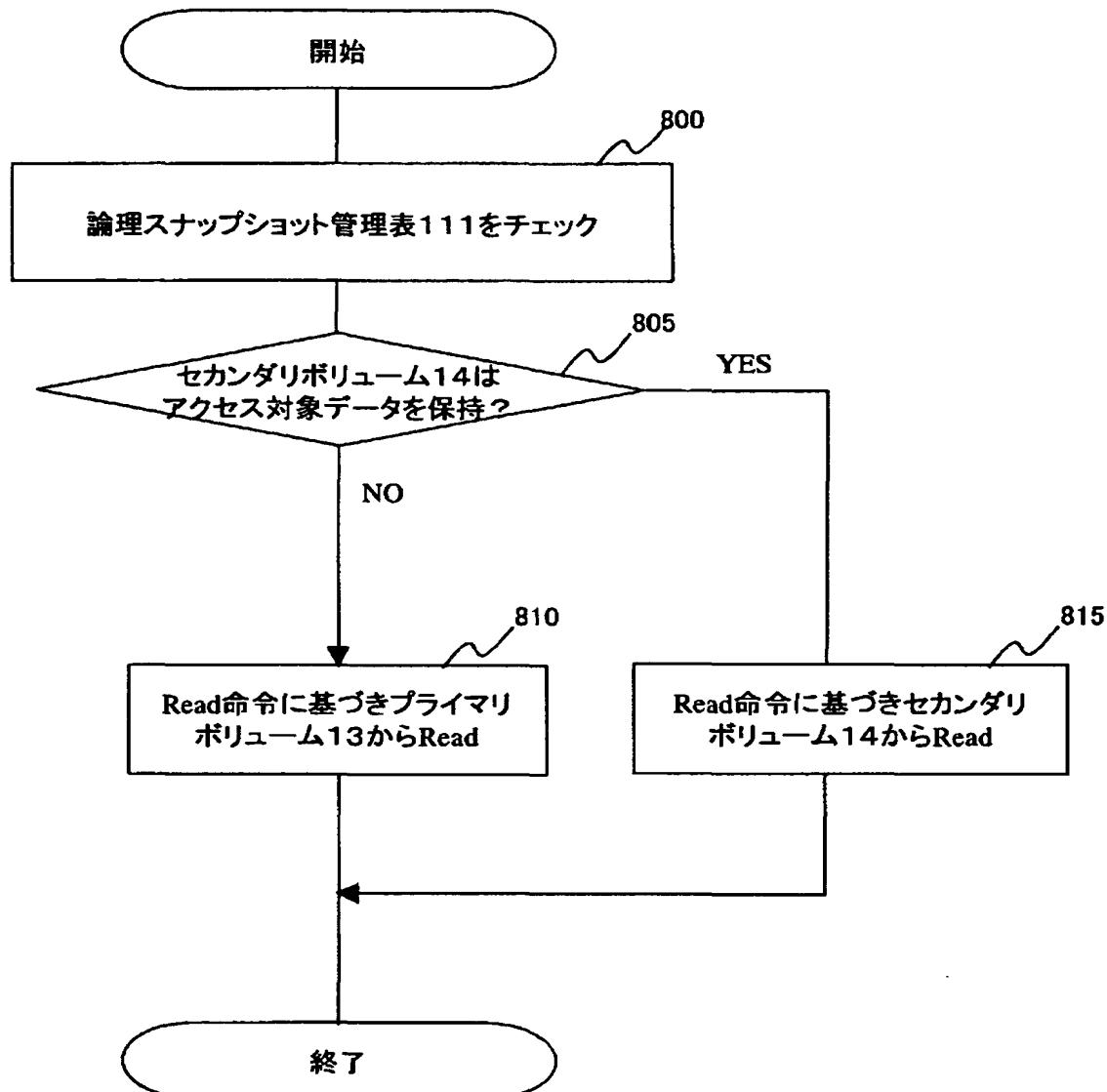
【図7】

図7



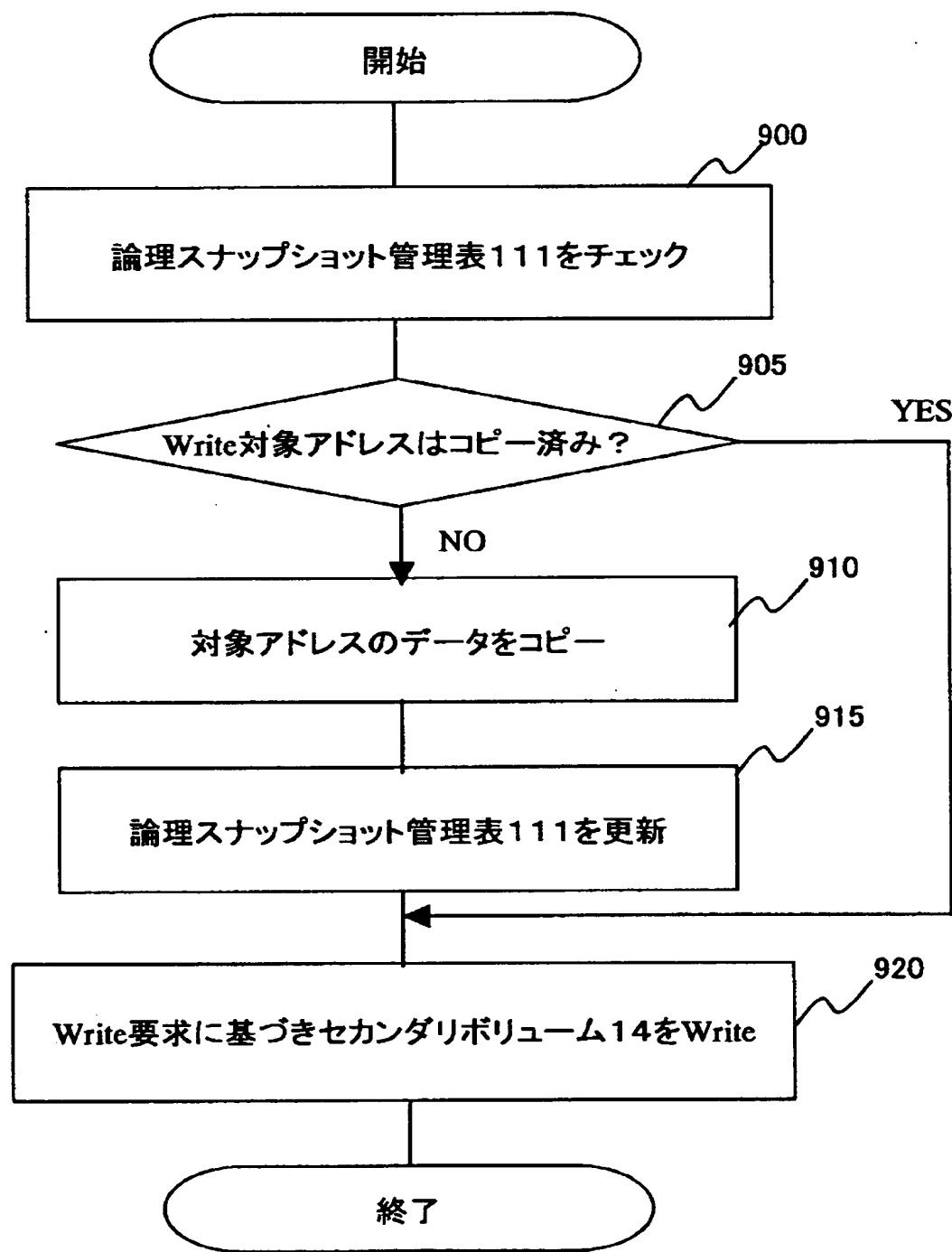
【図8】

図8



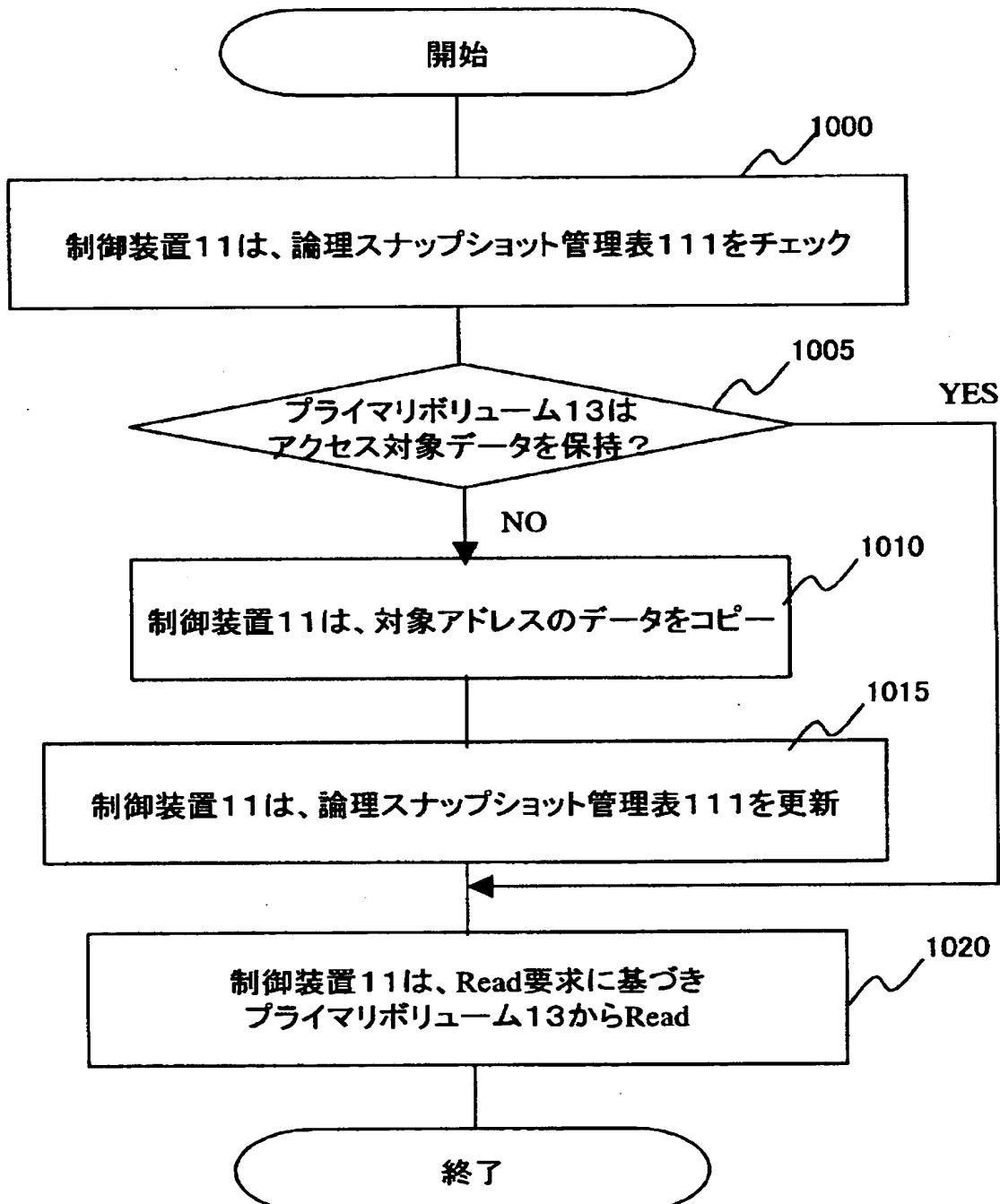
【図9】

図9



【図10】

図10



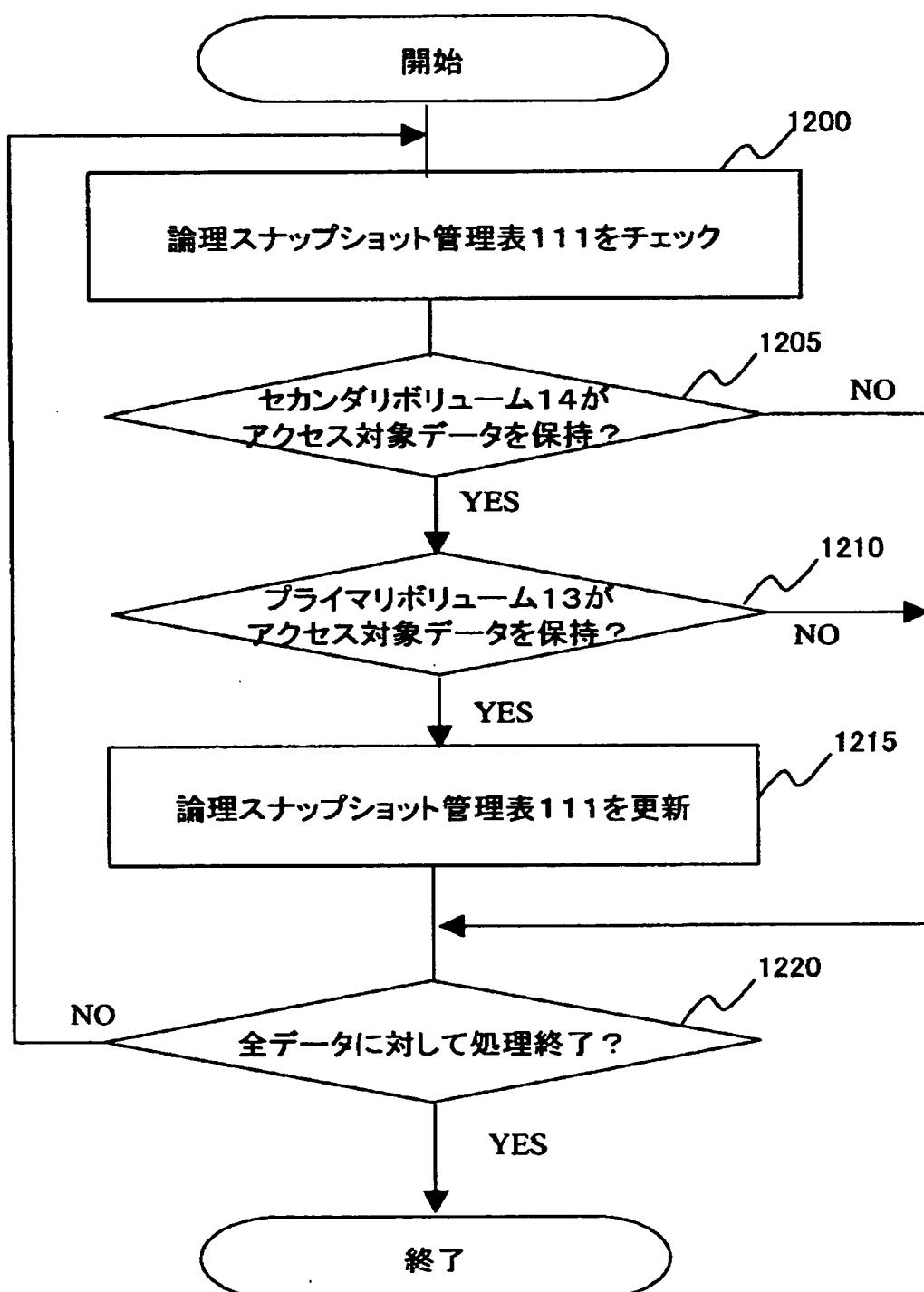
【図11】

図11

実行前の状態 (P, S)	Snap コマンド受領後の状態	Restore コマンド受領後の状態
(0, 0)	(0, 1)	(1, 0)
(0, 1)	(0, 1)	(0, 1)
(1, 0)	(1, 0)	(1, 0)
(1, 1)		N/A

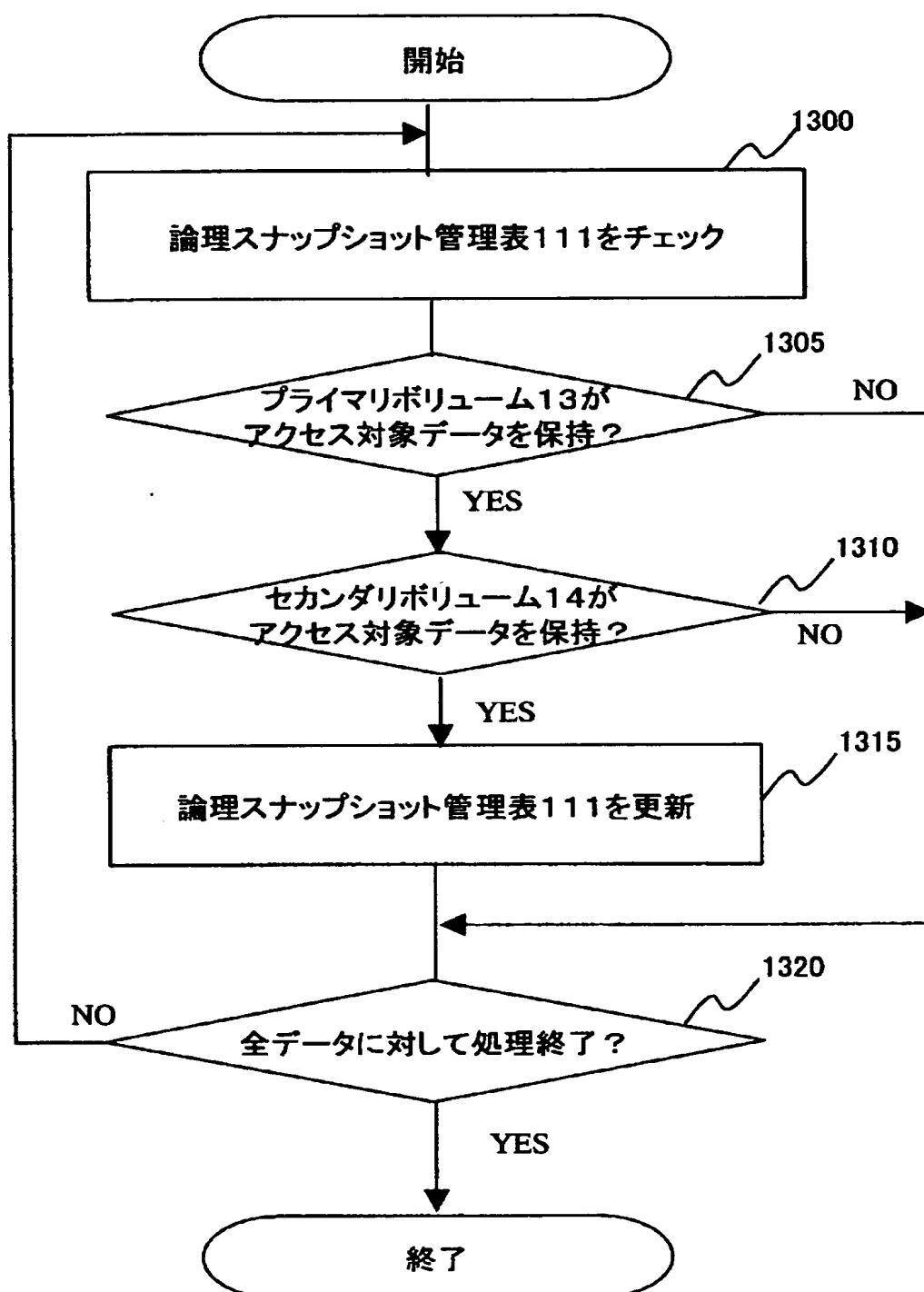
【図12】

図12



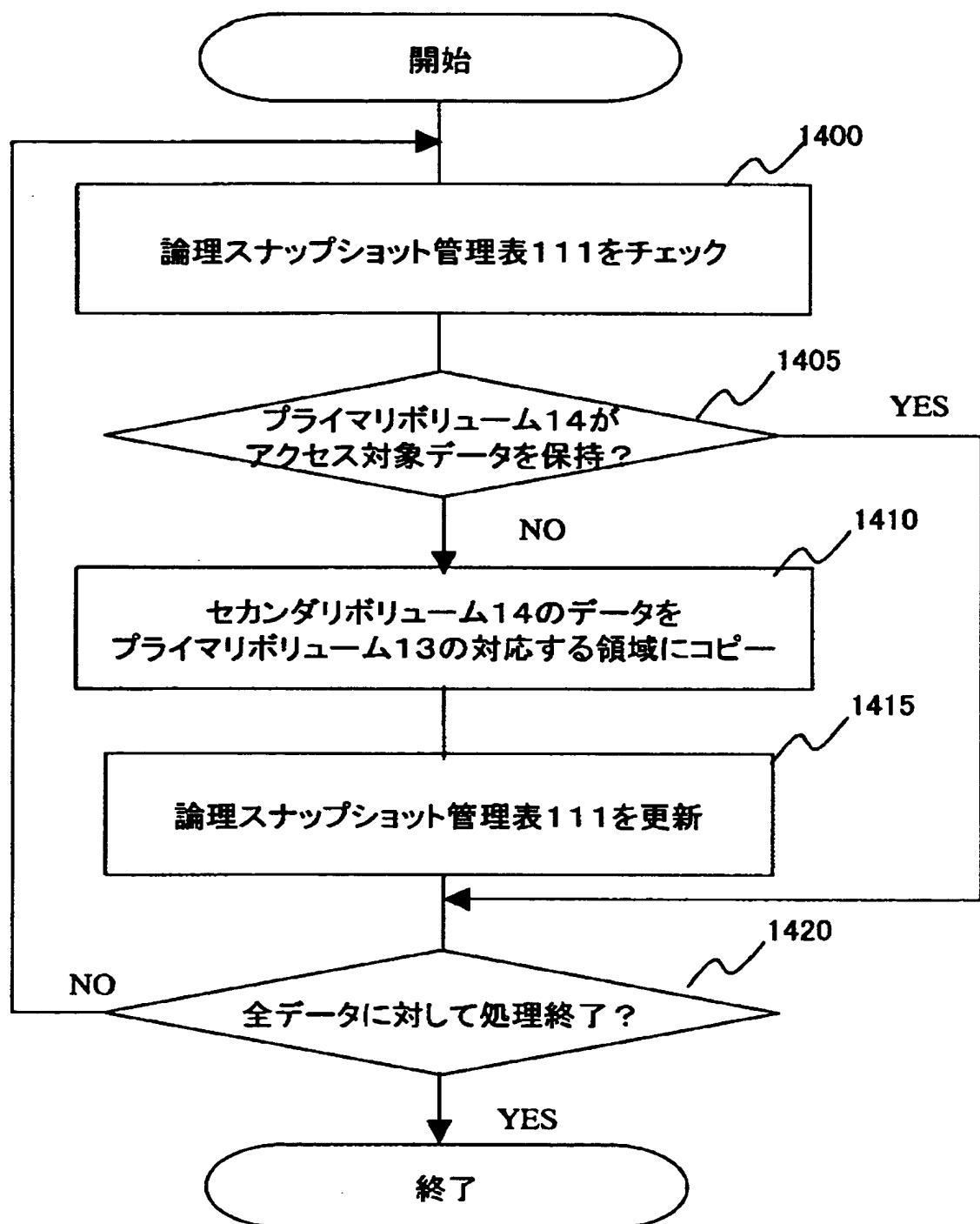
【図13】

図13



【図14】

14



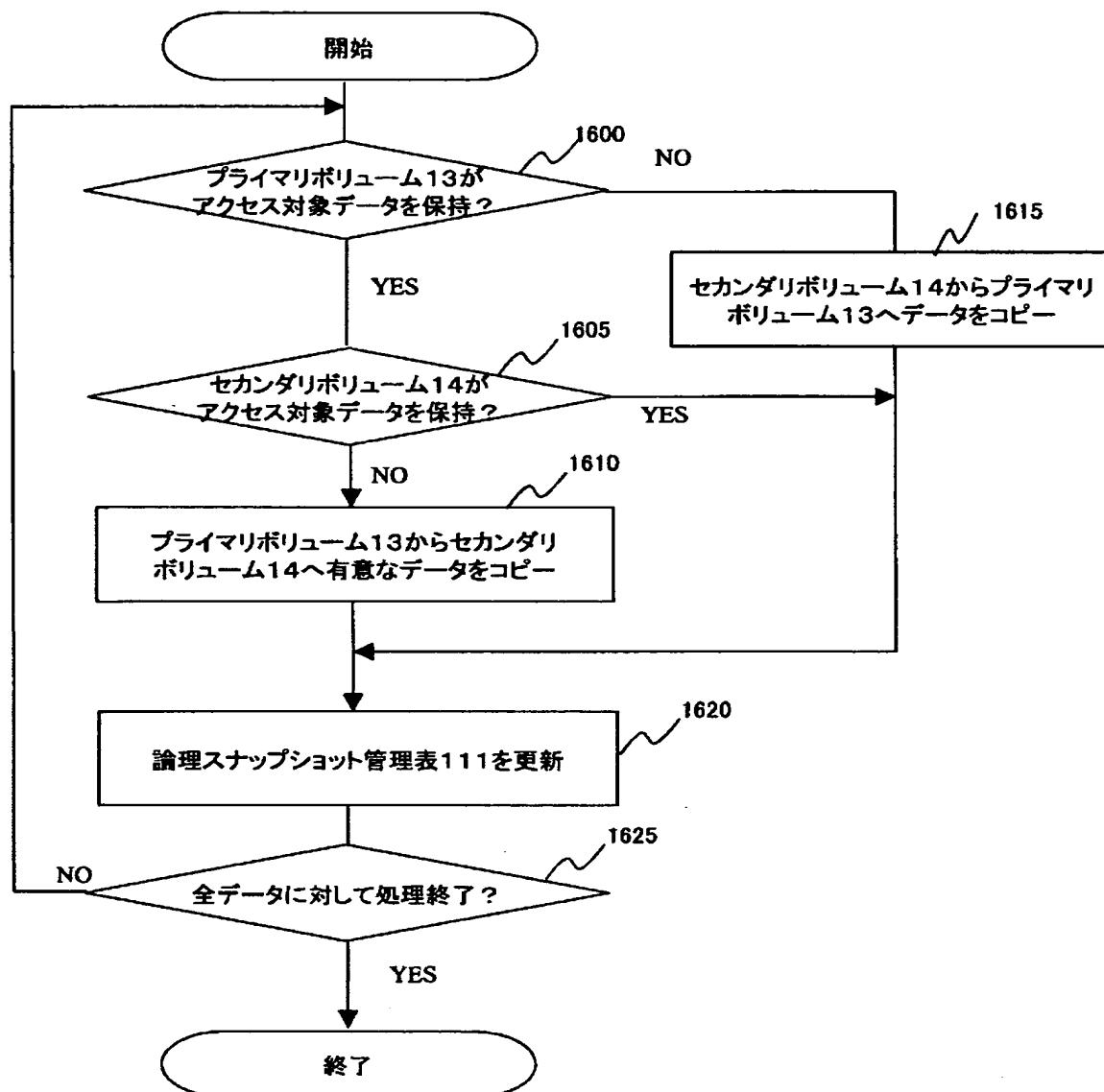
【図15】

図15

実行前の状態 (P, S)	Deleteコマンド受領後の状態
(0, 0)	(0, 0)
(0, 1)	(0, 1)
(1, 0)	(1, 0)
(1, 1)	N/A

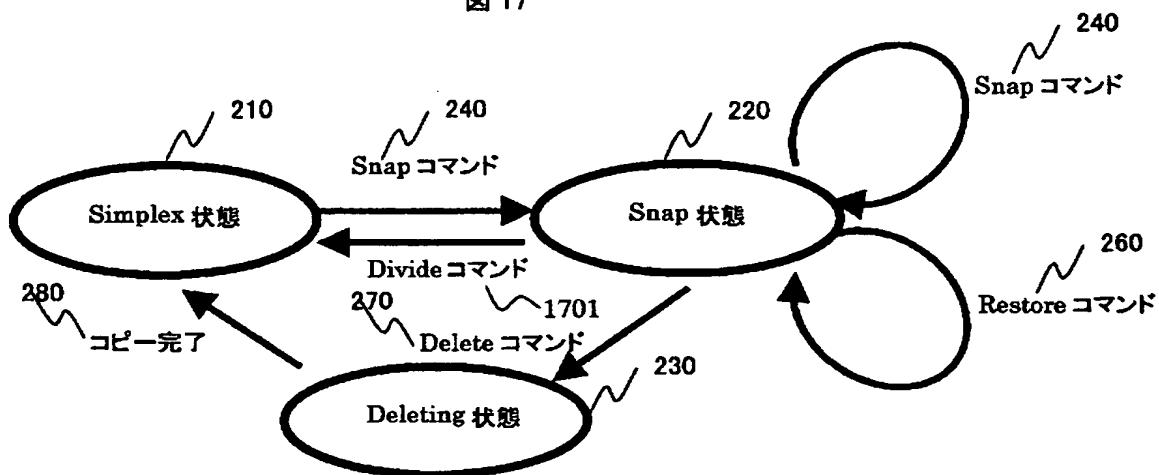
【図16】

図16



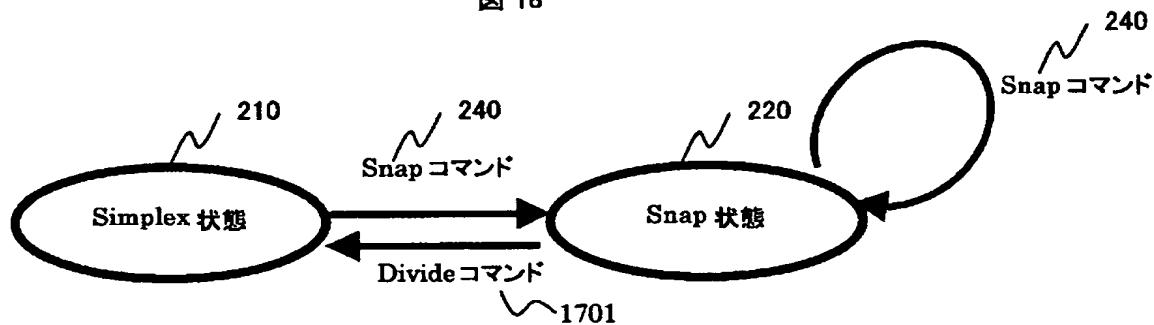
【図17】

図17



【図18】

図18



【書類名】 要約書

【要約】

ホストコンピュータの多様な使用環境においてもストレージあるいはホストに多大な負荷をかけずバックアップデータを取得するという利便性を持つストレージ装置およびその制御方法を提供する。

【課題】

特定時点におけるデータの凍結イメージにアクセスする従来技術はいくつか存在するが、そのいずれもユーザにとっての利便性が保証されているとは言えない。

【解決手段】

プライマリ、セカンダリボリュームを1対のペアとして制御し、アクセスすべきデータがいずれのボリュームに保持されているかを示す論理スナップショット管理表を設けることにより、論理的な凍結イメージへの即時アクセスを可能とする。

【選択図】 図1

認定・付加情報

特許出願の番号 特願2003-048483
受付番号 50300305670
書類名 特許願
担当官 第七担当上席 0096
作成日 平成15年 2月27日

＜認定情報・付加情報＞

【提出日】 平成15年 2月26日

次頁無

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地
氏 名 株式会社日立製作所